

Von visuellen Daten zu inkrementellen Wegbeschreibungen  
in dreidimensionalen Umgebungen:  
Das Modell eines kognitiven Agenten

Dissertation  
zur Erlangung des Grades  
des Doktors der Ingenieurwissenschaften  
der Technischen Fakultät  
der Universität des Saarlandes

von

Wolfgang Maaß

Saarbrücken  
1996

# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>1</b>
1.1	Problemstellung . . . . .	2
1.2	Einordnung der Arbeit . . . . .	2
1.3	Zielsetzung und Methodologie . . . . .	5
1.4	Gliederung . . . . .	7
<b>2</b>	<b>Theorien zur Raumkognition</b>	<b>9</b>
2.1	Zusammenhang zwischen Umgebungen und Situationen . . . . .	19
2.2	Räumliches Wissen in Situationen . . . . .	20
2.3	Raumkonzepte . . . . .	21
2.3.1	Räumliche Referenzsysteme . . . . .	24
2.3.2	Räumliche Relationen . . . . .	31
2.3.3	Wahl eines Referenzsystems im Kontext von Raumbeschreibungen . . . . .	34
2.3.4	Landmarkenwissen . . . . .	35
2.3.5	Routenwissen . . . . .	37
2.3.6	Konfigurationswissen . . . . .	38
2.3.7	Theorien zu Kognitiven Karten . . . . .	38
2.4	Aspekte des motorischen Verhaltens . . . . .	39
2.5	Interaktion von Sprache und räumlichem Wissens . . . . .	40
2.5.1	Raumdeixis . . . . .	43
2.5.2	Diskursbereich der Wegbeschreibungen . . . . .	43
2.5.2.1	Wegbeschreibungen als Diskurstyp . . . . .	44
2.5.2.2	Linearisierung durch reale und imaginierte Wanderungen . . . . .	45
2.5.2.3	Referentielle Ausdrücke in Wegbeschreibungen . . . . .	47
2.6	Modelle allgemeiner raumkognitiver Leistungen . . . . .	48
2.6.1	Das TOUR Modell . . . . .	48
2.6.2	Das Modell ELMER . . . . .	49
2.6.3	Das Modell von Elliott und Lesk . . . . .	50
2.6.4	Das Modell TRAVELLER . . . . .	50
2.6.5	Das Modell NAVIGATOR . . . . .	50
2.6.6	Das <i>Basic Agent</i> -Modell . . . . .	52
2.6.7	Das Modell SPAM . . . . .	53
2.6.8	Das Modell MERCATOR . . . . .	53

2.6.9	Das Modell von Yeap . . . . .	54
2.6.10	Zusammenfassung raumverarbeitender Modelle . . . . .	54
2.7	Modelle zur Generierung von Wegbeschreibungen . . . . .	57
2.7.1	Das formale Modell von Habel . . . . .	57
2.7.2	Das formale Modell von Höök und Karlgren . . . . .	58
2.7.3	Das formale Modell von Hoepfner et al. . . . .	59
2.7.4	Das formale Modell von Couclelis . . . . .	60
2.7.5	Das Modell von Rau und Schweitzer . . . . .	61
2.7.6	Das Berechnungsmodell KLEIST . . . . .	62
2.7.7	Das Berechnungsmodell CITYGUIDE . . . . .	63
2.7.8	Das Berechnungsmodell von Carstensen . . . . .	64
2.7.9	Das <i>Back Seat Driver</i> Modell . . . . .	66
2.7.10	Zusammenfassung der Wegbeschreibungsmodelle . . . . .	66
2.8	Zusammenfassung . . . . .	68
<b>3</b>	<b>Empirische Untersuchungen</b>	<b>71</b>
3.1	Vorbemerkungen . . . . .	71
3.2	Explorative Untersuchung in realen Umgebungen . . . . .	74
3.2.1	Arbeitsmodell . . . . .	74
3.2.2	Durchführung . . . . .	75
3.2.3	Ergebnis . . . . .	75
3.2.4	Diskussion . . . . .	81
3.3	Untersuchung in einer computeranimierten 3D-Umgebung . . . . .	82
3.3.1	Arbeitsmodell . . . . .	83
3.3.2	Durchführung . . . . .	83
3.3.2.1	Methode: Farbvariation . . . . .	85
3.3.2.2	Ergebnis: Farbvariation . . . . .	85
3.3.2.3	Methode: Höhenvariation . . . . .	87
3.3.2.4	Ergebnis: Höhenvariation . . . . .	87
3.3.3	Gesamtdiskussion . . . . .	88
<b>4</b>	<b>Formale Beschreibung des Agenten</b>	<b>91</b>
4.1	Modularer Aufbau des funktionalen Modells . . . . .	92
4.1.1	Das Steuerungsmodul . . . . .	93
4.1.2	Das Pfadsuchmodul . . . . .	94
4.1.3	Das Objektauswahlmodul . . . . .	96
4.1.4	Das Raumintegrationsmodul . . . . .	99
4.1.4.1	Räumliche Referenzsysteme . . . . .	99
4.1.4.2	Evaluation räumlicher Relationen . . . . .	100
4.1.4.3	Räumlicher Kurzzeitspeicher . . . . .	102
4.1.5	Das Navigationsmodul . . . . .	102
4.1.6	Das Sprachproduktionsmodul . . . . .	103

4.2	Funktionale Beschreibung . . . . .	103
4.2.1	Funktionale Beschreibung des Steuerungsmoduls . . . . .	105
4.2.1.1	Zeitrahen und Zeitbeschränkungen . . . . .	106
4.2.1.2	Dekomposition des Zeitrahmens $T_{proc}$ . . . . .	108
4.2.1.3	Aktivierung und Deaktivierung von Phasen . . . . .	110
4.2.1.4	Aktivierung und Evaluation von Ablaufplänen . . . . .	112
4.2.1.5	Funktionen des Steuerungsmoduls . . . . .	115
4.2.2	Funktionale Beschreibung des Pfadsuchmoduls . . . . .	117
4.2.2.1	Funktionen der inkrementellen heuristischen Pfadsuche . . . . .	118
4.2.2.2	Pfadabschnitte und Verzweigungspunkte . . . . .	121
4.2.2.3	Expansion von Verzweigungspunkten . . . . .	122
4.2.2.4	Zusammenfassung . . . . .	124
4.2.3	Funktionen des Objektauswahlmoduls . . . . .	125
4.2.3.1	Allgemeine Beschreibung der Auswahl visuell salienter Objekte . . . . .	125
4.2.3.2	Identifikation sichtbarer Objekte . . . . .	127
4.2.3.3	Abbildung der perspektivischen Sicht auf die Projektionsebene . . . . .	127
4.2.3.4	Berücksichtigung der räumlichen Fokussierung . . . . .	129
4.2.3.5	Abbildung der Projektions- auf die Lokationsebene . . . . .	130
4.2.3.6	Bestimmung visueller Salienzen bezüglich des rekursiven Merkmals „Farbe“ . . . . .	132
4.2.3.7	Bestimmung visueller Salienzen der linearen Merkmale „Höhe“ und „Breite“ . . . . .	137
4.2.3.8	Integration in globale Merkmalskarten . . . . .	138
4.2.3.9	Adaption der visuellen Salienzen an den räumlichen Fokusbereich . . . . .	140
4.2.3.10	Zusammenfassung . . . . .	141
4.2.4	Funktionen des Raumintegrationsmoduls . . . . .	141
4.2.4.1	Egozentrische und allozentrische räumliche Referenzsysteme . . . . .	141
4.2.4.2	Evaluation räumlicher Relationen . . . . .	150
4.2.4.3	Der räumliche Kurzzeitspeicher . . . . .	155
4.2.5	Funktionen des Navigationsmoduls . . . . .	166
4.2.6	Funktionen des Sprachproduktionsmoduls . . . . .	168
4.2.6.1	Generische Inhaltsstruktur inkrementeller Wegbeschreibungen . . . . .	169
4.2.6.2	Auswahl von Inhaltsstrukturen bezüglich einer Situation . . . . .	170
4.2.6.3	Beispiele der Verwendung räumlicher Konfigurationsbeschreibungen zur Generierung inkrementeller Wegbeschreibungen . . . . .	182
4.2.6.4	Der Präsentationsspeicher . . . . .	182
4.3	Implementation . . . . .	185
4.3.1	Prozesse der inkrementellen Pfadsuche . . . . .	186
4.3.2	Prozesse zur visuellen Objektauswahl . . . . .	190
4.3.3	Prozesse zur Verarbeitung räumlichen Wissens . . . . .	194
4.3.4	Prozesse zur Generierung natürlicher Sprache . . . . .	195
4.3.5	Beispiel einer automatisch generierten inkrementellen Wegbeschreibung . . . . .	196

<b>5</b>	<b>Ergebnisse und Ausblick</b>	<b>199</b>
5.1	Wissenschaftlicher Beitrag dieser Arbeit . . . . .	199
5.2	Einordnung in den technologischen Kontext . . . . .	201
5.2.1	Navigationshilfesystem in Realumgebungen . . . . .	203
5.2.2	Navigationshilfe in virtuellen Umgebungen . . . . .	208
5.3	Erweiterungsmöglichkeiten . . . . .	210

# Abbildungsverzeichnis

2.1	feature-integration-theory . . . . .	12
2.2	ebenen . . . . .	14
2.3	golledge-concepts . . . . .	22
2.4	is-a-concepts . . . . .	23
2.5	frame-of-ref . . . . .	25
2.6	f-o-r-deiktisch . . . . .	26
2.7	f-o-r-intrinsic . . . . .	26
2.8	f-o-r-extrinsic . . . . .	27
2.9	herrmann-ebene . . . . .	29
2.10	herrmann-deiktisch . . . . .	31
2.11	linearisierung . . . . .	46
2.12	navigator . . . . .	51
2.13	spat-struct-gopal . . . . .	51
2.14	habel-system . . . . .	58
2.15	coucelelis-system . . . . .	60
2.16	rau-system . . . . .	62
2.17	kleist-system . . . . .	63
2.18	carstensen-system . . . . .	65
3.1	karte-campus . . . . .	76
3.2	fragment-haeufig . . . . .	77
3.3	fragment-komb-haeuf . . . . .	78
3.4	zeitrahmen . . . . .	78
3.5	kat-pro-zeitrahmen . . . . .	80
3.6	kat-komb-pro-zrahmen . . . . .	80
3.7	exp-4-crossing . . . . .	84
3.8	color-intens . . . . .	86
3.9	right-wrong-side . . . . .	87
3.10	height-intens . . . . .	88
3.11	vgl-hoehe-farbe . . . . .	89
4.1	modules . . . . .	93
4.2	env2map-mapping . . . . .	94
4.3	pfad-ep . . . . .	96

4.4	spatial-module	99
4.5	anchoring	101
4.6	3d-campus-sgi	104
4.7	vollst-formales-modell	105
4.8	time-intervals	106
4.9	time-optimization	108
4.10	kontrollfluss-module	109
4.11	pfadintegrationsphase	111
4.12	objektintegrationsphase	112
4.13	WBgenerierungsphase	113
4.14	phases-flow	114
4.15	kontroll-modul	115
4.16	inkr-pfadsuche	118
4.17	corridor	119
4.18	entscheidungs-pkt	122
4.19	freier-platz	123
4.20	kreisverkehr	124
4.21	lineare+rek-merkmale	126
4.22	ausw-salienter-obj	128
4.23	projection	128
4.24	spatial-focus	129
4.25	focus-salience	130
4.26	sit2raster-locs	131
4.27	raster-error	131
4.28	cie-diagram	133
4.29	color-diffs	133
4.30	height-map-serra-sgi	138
4.31	width-map-serra-sgi	139
4.32	global-map-serra-sgi	139
4.33	sal-obj-1	140
4.34	dynamic-sit	142
4.35	prs-orientation	144
4.36	prim-ref-sys	144
4.37	srs-orientation	145
4.38	sek-ref-sys	145
4.39	sek-ref-sys2	146
4.40	vrs-orientation	147
4.41	virt-ref-sys	148
4.42	$f_{or} - tilted - A$	148
4.43	vrs-auto	149
4.44	@	152

4.45	right-of-deikt . . . . .	153
4.46	eval-intrinsic . . . . .	154
4.47	gen-rkb . . . . .	156
4.48	integr-path-elm . . . . .	157
4.49	integrate-path-segment . . . . .	157
4.50	topo-intrin-all . . . . .	158
4.51	path-intention-pattern . . . . .	160
4.52	proto-conf-intent . . . . .	160
4.53	select-srs . . . . .	162
4.54	integr-lm . . . . .	162
4.55	topo-intrin-all-lm . . . . .	164
4.56	loc-of-lms . . . . .	165
4.57	appl-of-lm . . . . .	166
4.58	anchoring-prs-in-nav . . . . .	167
4.59	animation-seq-move . . . . .	167
4.60	animation-seq-rot . . . . .	168
4.61	message-structure . . . . .	170
4.62	mesg-structure-partial . . . . .	173
4.63	lm-not-refo . . . . .	176
4.64	desc-lm . . . . .	177
4.65	pagm . . . . .	178
4.66	inh-kat-komb . . . . .	183
4.67	gui-moses . . . . .	186
4.68	gui-campus . . . . .	188
4.69	fuss-increm-2jams . . . . .	189
4.70	lf1 . . . . .	189
4.71	lf10 . . . . .	190
4.72	gui-moses-view . . . . .	190
4.73	gui-moses-speak . . . . .	195
4.74	compl-descr-foot.pic . . . . .	196
4.75	compl-descr-car.pic . . . . .	197

*“What you going to do with yourself, Ed?” I asked. “I don’t know,” he said, “I just going along. I dig life’.” He repeated it, following Dean’s line. He had no direction.*

*(Jack Kerouac: On the Road, 1957)*

# Kapitel 1

## Einführung

Stellen Sie sich vor, Sie kommen in Saarbrücken am Hauptbahnhof an und möchten zum St. Johanner Markt, kennen sich aber überhaupt nicht aus. Eine typische Lösungsstrategie ist es, wenn Sie jemanden nach dem Weg dorthin fragen. Unterwegs werden Sie in den meisten Fällen unsicher und fragen erneut eine Person nach dem Weg. In einem solchen Szenario wünschen Sie sich einen Ortskundigen an ihrer Seite, der Sie auf dem Weg zum Ziel begleitet und unterwegs zu geeigneten Zeitpunkten die richtige wegrelevante Information mitteilt, wie zum Beispiel: „An der nächsten Kreuzung müssen wir hinter dem großen Haus auf der linken Seite links ab.“ Eine soeben beschriebene Situation ist mit einer solchen zwischen Fahrer und Beifahrer in einem Auto vergleichbar. Die Aufgabe des Auskunftgebenden, im weiteren kurz „Sprecher“ genannt, besteht darin dem Fragenden, kurz „Rezipient“ genannt, in geeigneter Weise zu richtigen Zeitpunkten Information über den Weg mitzuteilen. Der Sprecher bezieht sich in seinen Beschreibungen vor allem auf visuell wahrnehmbare Entitäten der Situation, unter der Voraussetzung, daß der Rezipient diese leicht erkennen kann.

Wegbeschreibungen lassen sich prinzipiell in *vollständige* und *inkrementelle* unterteilen. Erstere werden gegeben, wenn der Sprecher einen Weg komplett beschreibt, ohne daß der Rezipient die referenzierten Elemente der Beschreibung direkt in der Situation verifizieren kann. Inkrementelle Wegbeschreibungen hingegen werden schrittweise gegeben. Durch sie werden Abschnitte eines Weges und geeignete Objekte beschrieben, die für den Rezipienten zum Identifizieren des geeigneten Weges in der jeweils aktuell perzipierten Situation relevant sind. Dadurch hat der Rezipient die Möglichkeit, die Beschreibung augenblicklich in der Situation zu verifizieren. Inkrementelle Wegbeschreibungen stellen folglich geringere Anforderungen an die Gedächtnisleistungen des Rezipienten, jedoch höhere an die Adaptation der Beschreibung an die Situation. Der Rezipient ist nicht, wie bei vollständigen Wegbeschreibungen, alleine auf die sprachliche Beschreibung angewiesen, sondern greift zusätzlich auf Information zurück, die durch Wahrnehmung aufgenommen worden ist. Folglich können inkrementelle Wegbeschreibungen nur im Verbund mit perzipierter Information kommuniziert werden.

## 1.1 Problemstellung

Um sich adaptiv und nutzbringend in seiner Alltagsumgebung verhalten zu können, ist es für den Menschen entscheidend, aus der Umgebung die Information herauszufiltern, die bezüglich seiner Intentionen wichtig erscheint. Von einer internen Perspektive aus gesehen, bedingen Kapazitätsbeschränkungen des menschlichen kognitiven Systems, daß er sich bei der Auswahl stark einschränken muß. Neben diesen internen Beschränkungen sind externe Beschränkungen zu beachten, die durch die Umgebung oder das Verhalten des Menschen in der Umgebung vorgegeben sind. Beispielsweise gibt die Bewegung eines Sprechers relativ zur nächsten Kreuzung vor, inwieweit sein Verhalten zeitlich und räumlich eingeschränkt ist. Aus dem jeweiligen Sachverhalt leitet der Sprecher die Beschränkungen ab, verinnerlicht diese und macht sie zur Grundlage seiner Verhaltensfindung. Diese Beschränkungen bedingen, daß der Mensch in möglichst *effizienter* Weise Information auswählt, um in *effektiver* Weise agieren zu können.

Ein solches Zusammenspiel von Umgebung und menschlichem Akteur wird in dieser Arbeit am Beispiel inkrementeller Wegbeschreibungen untersucht und formalisiert. In diesem Kontext läßt sich in besonderer Weise untersuchen, wie sprachliche Beschreibungen vom Sprecher an dynamisch veränderliche Umgebungen adaptiert werden. Die Prozesse, die zur Generierung inkrementeller Wegbeschreibungen vonnöten sind, lassen sich in vier Teile zerlegen. Zum einen analysiert der Sprecher seine visuell sichtbare Umgebung und extrahiert Entitäten, die ihm „ins Auge fallen“, d.h., die eine besondere *visuelle Salienz* besitzen. Um einen Weg beschreiben zu können, muß der Sprecher zweitens den Weg bestimmen. Entweder besitzt er ausreichendes räumliches Wissen über die Umgebung oder er greift auf externe Hilfsmittel, wie beispielsweise *Straßenkarten* zurück. In dieser Arbeit wird der letzte Fall betrachtet. Die aus der Umgebung und der Karte extrahierte Information wird vom Sprecher drittens dazu verwendet, eine *räumliche Vorstellung* seiner Umgebung aufzubauen und dynamisch an Veränderungen anzupassen. Im Zusammenhang mit Intentionen und Rezipienteneigenschaften, verwendet der Sprecher diese Vorstellungen, um viertens Beschreibungen des Weges schrittweise zu *generieren* und zum richtigen Zeitpunkt zu *präsentieren*. Der Mensch ist in der Lage, die kognitiven Leistungen, die mit diesen Einzelphänomenen verbunden sind, miteinander zu koordinieren und an interne und externe Beschränkungen zu adaptieren. In dem hier vorgestellten formalen Modell eines kognitiven Agenten werden ebensolche Beschränkungen bei der Handlungsfindung- und ausführung berücksichtigt.

In Tabelle 1.1 sind zum Vergleich eine vollständige und eine inkrementelle Wegbeschreibung gegenübergestellt. Sie beschreiben jeweils denselben Pfad vom Gebäude der Informatik zur Hauptbibliothek der Universität des Saarlandes. Auffällig ist, daß die vollständige im Vergleich zur inkrementellen Wegbeschreibung wesentlich kürzer und ungenauer ist. Ferner werden vielmehr große Gebäude referenziert.

## 1.2 Einordnung der Arbeit

Dieses derart strukturierte Phänomen liegt im Spannungsfeld zwischen *visueller Wahrnehmung*, *Verarbeitung räumlichen Wissens* und *Sprachproduktion*. In dieser Arbeit wird das Phänomen aus zwei Perspektiven betrachtet. Einerseits werden durch eine *kognitionswissenschaftliche Perspektive* Fragestellungen und Konzeptualisierungen fokussiert, die im Hinblick auf Simulationen von Fähigkeiten und Verhalten des Menschen eine Rolle spielen. Die verschiedenen Disziplinen, wie Psychologie, Anthro-Geographie, Neurobiologie, Psycho- und Computerlinguistik und die der Künstlichen Intelligenz geben Hinweise darüber, welche informationsverarbeitende Mechanismen involviert sind. Aus einer *Ingenieursperspektive* heraus ist andererseits die Verwendbarkeit eines Systems zu untersuchen, welches in der Lage ist, inkrementelle Wegbeschreibungen zu ge-

<p>Also, dann gehe unter der Unterführung durch, an der Mathematik vorbei,</p> <p>auf der rechten Seite kommt dann ein freier Platz, gehe über den Platz geradeaus die Treppe hoch,</p> <p>dann kommst du an eine Straßensache, wo du so links gehst,</p> <p>dann siehst du den EDEKA-Markt,</p> <p>dann gehts rechts an diesem EDEKA vorbei,</p> <p>dann ist da auch die Bayerische Vereinsbank,</p> <p>dann gehst du hoch,</p> <p>dann links abbiegen,</p> <p>dann bei der Post rechts abbiegen und dann links die Treppen hoch,</p> <p>dann bist du auf einem großen freien Platz, dort nach rechts,</p> <p>und dann sieht man auch schon den hohen Büchereiturm,</p> <p>Aber bei Bock und Seip solltest du am besten nochmal nachfragen</p>	<p>Hier jetzt geradeaus, die Treppe runter, über die Straße dann da die Treppe wieder hoch.</p> <p>Geradeaus. Weiter geradeaus, unter der Überdachung durch. Jetzt gleich hier vorne rechts.</p> <p>Hier jetzt wieder ein Stück geradeaus ... an dem Gebäude. Jetzt kommen wieder Treppen ... hoch.</p> <p>Hier über den Weg.</p> <p>Und dann die großen Treppen da vorne wieder hoch.</p> <p>Geradeaus hier. Und da vorne bei der Wegkreuzung vor dem Glaskontainer oder Altpapierkontainer ... dann links ... OK, jetzt links hier.</p> <p>Hier geradeaus ein Stück und oben an der Kreuzung dann wieder rechts. So jetzt hier rechts.</p> <p>Geradeaus. Treppe hoch. Und hier vorne jetzt an der Kreuzung links.</p> <p>Hier jetzt geradeaus. Und vor dem Musikbau da vorne auf der linken Seite gehts rechts die Treppe hoch. Also zwischen Post und Musikgebäude. So da jetzt rechts rüber, da wo die Leute herkommen.</p> <p>Hier jetzt links die Treppe hoch.</p> <p>Jetzt, hier rechts.</p> <p>Jetzt einfach geradeaus. Und da vorne ist auch schon die Bibliothek. Hier jetzt einfach über die Straße rüber ... über den Zebrastreifen und da wo Fahrräder und die Motorräder stehen, da ist der Eingang. Da gehts rein.</p>
---	---

Tabelle 1.1: Beispiele für eine vollständige und eine inkrementelle Wegbeschreibung

nerieren. Navigationshilfesysteme werden bereits punktuell im Automobilverkehr eingesetzt. Sie besitzen jedoch einen viel größeren Einsatzbereich, wie nicht zuletzt zu Rehabilitationszwecken bei neurophysiologisch oder perzeptuell geschädigten Personen. Beide Perspektiven lassen sich unter der Prämisse zusammenfassen, daß ein Navigationshilfesystem um so nutzbringender ist, je mehr es sich wie ein Mensch verhält. Dies aber kann nur dann verwirklicht werden, wenn man die Prozesse und Repräsentationen, die das Verhalten des Menschen determinieren, besser versteht.

Die vorliegende Arbeit ist im Kontext der Künstlichen Intelligenz (KI) Forschung erstellt worden. Eines der Arbeitsgebiete befaßt sich mit der Interaktion von visueller Wahrnehmung, räumlichem Wissen und Sprache. Besonders in Deutschland ist die Forschung in dieser Richtung sowohl aus kognitionswissenschaftlicher als auch ingenieurwissenschaftlicher Sicht vorangetrieben worden. Bereits mit dem System HAM-RPM (vgl. [WJH78]) und nachfolgend mit SWYSS (vgl. [HS84]) wurde ein einfaches Modell der natürlichsprachlichen Beschreibung statischer räumlicher Infor-

mation vorgestellt, wobei insbesondere auch Aspekte der Vagheit sprachlicher Raumbeschreibungen ihre Berücksichtigung fanden. Während in SWYSS eine Seitenansicht eines Zimmers als Ausgangsszene behandelt wurde, stand in den Systemen GEOSYS (vgl. [FKR84]), LANDSCAN (vgl. [BJKZ85]) und LEI (vgl. [CMC94]) der Zugriff auf geographische Information und die damit verbundenen räumlichen Beziehungen im Vordergrund. Die Aufgabenstellung in den Systemen WINTOP (vgl. [OMT94]) und VIENA (vgl. [WC94]) ist es, geometrische Szenenmodelle in einem Graphiksystem durch sprachliche Interaktion inkrementell zu verfeinern. Neben rein softwarebasierten Ansätzen wird in einigen Arbeiten untersucht, wie die Integration von visuellen Daten und Sprache zur Steuerung von Robotern verwendet werden kann (vgl. [LFMM94], [HMRS95], [SGH<sup>+</sup>94]). Das Projekt LILOG-R ([HR91], [Pri93]) untersuchte vor allem Repräsentationen, die das *Verstehen* von Lokalisierungsausdrücken bezüglich Situationen, zu denen der Rezipient keinen unmittelbaren perzeptuellen Zugang besaß, modellieren. Ein Überblick zu den aktuellen Forschungsaktivitäten im Bereich der Integration von perzeptuellen und sprachlichen Fähigkeiten findet sich in verschiedenen Aufsatzsammlungen (vgl. [McK94b], [McK94a]).

Im Kontext der Integration von *bildverstehenden* und *sprachverstehenden Systemen* ist die Domäne der Wegbeschreibungen, aufgrund der Abgegrenztheit des Phänomenbereichs, aus linguistischer, psychologischer und KI-Sicht bezüglich des dynamischen Zusammenspiels von visueller Perzeption, Raumkognition und Sprache häufig untersucht worden. In linguistischen Arbeiten sind vor allem syntaktische und semantische Aspekte von sprachlichen Wegbeschreibungen analysiert worden (z.B. [Kle82], [WR82], [Hab87], [MMP<sup>+</sup>88], [HCR90]). Psychologisch motivierte Arbeiten haben anhand von Wegbeschreibungen besonders räumliches Wissen untersucht (z.B. [HM71], [DS73], [Tve92]). Beide Aspekte spielen in algorithmischen Modellen eine wesentliche Rolle (z.B. [Kui78], [Dav86], [GKS89], [LZ89]).

Die hier vorgestellte Arbeit ist im Kontext der Projekte VITRA<sup>1</sup> und REAL<sup>2</sup> entstanden. In VITRA ist es gelungen, erste Erkenntnisse über die Durchschaltung visueller Realdaten zu sprachlichen Beschreibungen zu gewinnen ([HSA<sup>+</sup>89]). Bezüglich der Domäne des Fußballspiels ist im Rahmen von VITRA ein System namens SOCCER entwickelt und implementiert worden, welches visuelle Daten von realen Fußballszenen verbal beschreibt (vgl. [HRA89], [HW94], [Sch94]). Weiterhin wird in einer Konstruktionsdomäne untersucht, wie ein Roboter mit zwei Effektoren natürlichsprachlich gesteuert werden kann (vgl. [SL96]).

In Zusammenarbeit mit Forschern der *Visuellen Datenverarbeitungsgruppe* am *Fraunhofer Institut (IITB)* in Karlsruhe sind im Rahmen des Sonderforschungsbereichs 314 „Wissensbasierte Systeme“ Techniken untersucht worden, mit denen Videoaufnahmen von Fußballspiel- und Verkehrsszenen analysiert und verbal beschrieben werden können. Als Schnittstelle zwischen visuellen Daten und sprachlichen Beschreibungen werden *geometrische Szenenbeschreibungen* verwendet, die zeitlich ausgedehnte Szenen über die Festlegung von Lokationen relativ zu einem Koordinatensystem integrieren (vgl. [Neu86]). Auf der Basis der geometrischen Szenenbeschreibung wird die statische und dynamische, räumliche Struktur der Szene analysiert. Einerseits werden dazu statische, räumliche Lagebeziehungen zwischen Objekten ([And88]) und andererseits Ereignisse erkannt (vgl. [HR88], [Her96]). Neben diesen rein räumlichen Analysen, können situationsbezogene Intentionen abgeleitet werden (vgl. [RS92]). Räumliche Beschreibungen erzeugen beim Rezipienten räumliche Vorstellungen. Perzipiert der Rezipient den beschriebenen Sachverhalt, so kann er/sie diese räumliche Vorstellung sofort verifizieren. In Situationen, wie beispielsweise bei einer Radioreportage eines Fußballspiels, ist dies nicht möglich. In einem solchen Fall ist eine solche Beschreibung zu generieren, die es dem Rezipienten erlaubt, eine geeignete

<sup>1</sup>VITRA steht für “visual translator”, welches das Teilprojekt N2 des SFB 314 (Wissensbasierte Systeme) war (vgl. [?], [HBG<sup>+</sup>96], [Wah94]).

<sup>2</sup>REAL steht für “Ressourcenadaptive Objektlokalisierung”, welches das Teilprojekt A4 des SFB378 (Ressourcenadaptive kognitive Prozesse) ist (vgl. [TLS<sup>+</sup>95]).

mentale räumliche Repräsentation der Szene aufzubauen. Bevor der Sprecher eine Beschreibung präsentiert, antizipiert er durch Konstruktion eines „mentalenes Bildes“ die Vorstellung beim Rezipienten, die durch die Beschreibung erzeugt wird, und paßt seine Beschreibung gegebenenfalls an (vgl. [WJH78], [Sch94]). Ein Berechnungsmodell einer solchen *Antizipationsrückkopplungsschleife* ist in SOCCER integriert (vgl. [Sch94],[SS93], [BS95]). Im Kontext von VITRA ist mit dem System CITYGUIDE ein erstes Modell für die Generierung von vollständigen Wegbeschreibungen entwickelt worden (vgl. [Mül88]). Als Weiterentwicklung der in VITRA entwickelten Konzepte entstand die Fragestellung nach einem Modell zur Generierung inkrementeller Wegbeschreibungen. Die Zielsetzung und die verwendete Methodologie sollen nun kurz vorgestellt werden.

### 1.3 Zielsetzung und Methodologie

Die hier vorgestellte Arbeit hat zum Ziel, am Beispiel inkrementeller Wegbeschreibungen zu untersuchen, wie zeitlich-räumliche Beschränkungen zwischen einem Sprecher und einer räumlichen Umgebung das sprachliche Verhalten bezüglich Raumbeschreibungen beeinflußt. Inkrementelle Wegbeschreibungen sind aufgrund der Eigenbewegung des Sprechers an die Beschränkungen und Eigenschaften der Umgebung und davon abgeleiteten Ressourcenbeschränkungen anzupassen. Zentrales Thema dieser Arbeit ist es, jene Komponenten zu formalisieren und zu integrieren, die notwendig sind, um in einer gegebenen Situation inkrementelle Wegbeschreibungen zu generieren. Ausgehend von der Annahme, daß der Sprecher und der Rezipient über die Umgebung keinerlei räumliche Vorstellung besitzen, werden im Kontext inkrementeller Wegbeschreibungen folgende Kernpunkte behandelt:

#### 1. *Zeitabhängige Generierung* inkrementeller Wegbeschreibungen

Wie die empirischen Untersuchungen ferner zeigen, wird eine inkrementelle Wegbeschreibung an zeitliche und räumliche Ressourcenbeschränkungen adaptiert, die durch die Bewegung des Sprechers relativ zur Umgebung gegeben sind. Die Komplexität einer Beschreibung hängt von dem Zeitraum ab, den der Sprecher bis zum Erreichen des nächsten Verzweigungspunktes zur Verfügung hat. Entsprechend dieser Beschränkungen werden geeignete Inhaltsstrukturen ausgewählt, Oberflächenstrukturen transformiert und zu geeigneten Zeitpunkten präsentiert.

#### 2. Auswahl *visuell salienter Objekte* in simulierten 3D-Umgebungen

Die *visuelle Salienz* eines Objektes hängt von seinen visuellen Merkmalen und dem Kontext, der durch die gesamte Situation gegeben ist, ab. Visuelle Salienzen lassen sich bezüglich einzelner Merkmalsklassen bestimmen, die, relativ zur Umgebung gewichtet, zu einem Salienzwert eines Objektes integriert werden.

#### 3. Inkrementelle heuristische Suche von *Pfaden* in topographischen Netzen

Ausgehend davon, daß der Sprecher kein räumliches Wissen bezüglich der Umgebung besitzt, wird eine topographieerhaltende Kartenrepräsentation verwendet, um geeignete Pfadelemente (Pfadabschnitte und Kreuzungen) inkrementell unter Einbeziehung von Wissen zu bestimmen.

#### 4. Etablierung geeigneter *räumlicher Referenzsysteme*

In einer Umgebung benötigt ein Sprecher eine Orientierung, auf dessen Basis er die räumliche Lage von Objekten zu ihm selbst oder relativ zwischen Objekten bestimmen kann. Zu diesem Zweck sind verschiedene räumliche Referenzsysteme zu etablieren, die durch

die Bewegung oder Blickrichtung des Sprechers bzw. durch intrinsische Eigenschaften der Objekte gegeben sind.

5. Etablierung einer *minimalen* Anzahl *räumlicher Relationen*

Unter der Annahme, daß der Mensch aus Ökonomie- und Effizienzgründen versucht, eine möglichst geringe Anzahl an räumlichem Wissen aufzubauen, wird analysiert, zwischen welchen Objekten räumliche Relationen etabliert werden müssen, um Wegbeschreibungen generieren zu können. Hierzu werden gradierte räumliche Relationen verwendet.

6. Schemabasierte Ableitung *pfadgeleiteter Intentionen*

Die globale Intention des Sprechers ist es, den Rezipienten zur Ziellokation zu führen. Auf lokaler Ebene muß der Sprecher evaluieren, in welche Richtung der Pfad an einer Verzweigungsmöglichkeit verfolgt werden soll. Die Lage der Pfadelemente zueinander beeinflusst, welche Intention der Sprecher für die nachfolgend auszuführende Bewegung und somit für die Beschreibung zugrunde legt. Solche Intentionen werden als *pfadgeleitet* bezeichnet.

7. Situationsadaptive Auswahl sprachlicher *Inhaltsstrukturen*

Wie die in Kapitel 3 vorgestellten empirischen Untersuchungen zeigen, werden deutschsprachige inkrementelle Wegbeschreibungen meist durch schematische Kombination einzelner inhaltlicher Teile aufgebaut. Diese Teile (Inhaltskategorien) beziehen sich auf die Beschreibung einer Lokation, einer Richtung, eines Zeitpunktes oder einer Aktion und werden zu *Inhaltsstrukturen* zusammengesetzt.

8. Schemabasierte Auswahl räumlicher *Präpositionen* und *Adverbien*

Räumliche Information wird in Wegbeschreibungen vor allem durch Präpositionen und Adverbien ausgedrückt. Es wird auf der Basis des dynamisch etablierten räumlichen Wissens ein Verfahren zu Lexikalisierung vorgestellt.

Da bisher kein Korpus inkrementeller Wegbeschreibungen frei verfügbar ist, mußte als Teil dieser Arbeit ein entsprechender Korpus durch eine Felduntersuchung aufgebaut werden.<sup>3</sup> Aus diesem Korpus sind inhaltliche Strukturen und zeitliche Abhängigkeiten abgeleitet worden. In einer zweiten empirischen Untersuchung ist bezüglich einer computeranimierten 3D-Umgebung analysiert worden, in welcher Weise die Variation visueller Merkmale und pfadbasierter Intentionen die Auswahl von Landmarken, die meist als Referenzobjekte verwendet werden, beeinflusst.

Durch Integration raumkognitiver Theorien und solchen aus empirischen Untersuchungen, ist es das Ziel, durch Synthese ein formales Modell abzuleiten. Das formale Modell konstituiert einen Agenten, welcher sich durch ein dreidimensionales Modell des Campus der Universität des Saarlandes bewegt und dabei inkrementelle Wegbeschreibungen generiert. Um dieses ohnehin komplexe Problem bewältigen zu können, wird in dieser Arbeit auf eine dialogische Situation verzichtet. Weiterhin wird eine maximale Kooperationsbereitschaft des Rezipienten vorausgesetzt. Als Bedingung an die Formalisierung werden nur solche Repräsentationen und Prozesse verwendet, die unbedingt zur Generierung des Verhaltens notwendig sind. Der durch das Modell spezifizierte Agent soll verschiedenen Anforderungen entsprechen. So soll der Agent sich an situative Änderungen *adaptieren*. Dies bezieht sich vor allem auf solche Änderungen, die durch die Eigenbewegung des Agenten induziert werden. Lokationen von Objekten in einer Situation ändern sich dabei relativ zur Lokation des Agenten in der Zeit. Daraus resultierende, *externe* Ressourcenbeschränkungen müssen intern repräsentiert und für die Adaption des Verhaltens

<sup>3</sup>Am LIMSI in Orsay ist in einem Fahrer-Beifahrer-Szenario ein 64-stündiger Korpus französischer inkrementeller Wegbeschreibungen aufgebaut worden (vgl. [BD95]). Da es sich hierbei um eine Auftragsarbeit für einen Automobilkonzern handelt, ist der Korpus nicht frei verfügbar.

verwendet werden. Weiterhin soll das formale Modell nicht auf eine, sondern auf eine Klasse von Domänen anwendbar sein, weswegen das Modell *flexibel* sein muß. Die hier diskutierte Domänenklasse ist die *städtischer Umgebungen*, welche durch die Objektklassen *Straßen*, *Kreuzungen* und *Gebäude* strukturiert sind. Eine solche Einschränkung auf eine Domänenklasse gewährleistet eine weitgehende Homogenität in der Verwendung der Objektklassen. Dennoch ist das funktionale Modell derart konzipiert, daß es weitgehend generisch ist, so daß es mit geringem Aufwand an andere Domänen adaptiert werden kann.

Fokussiert auf die Verhaltensgenerierung sind zwei weitere Forderungen zu stellen, die sich wesentlich auf agentenspezifische Eigenschaften beziehen. Es ist zu fordern, daß die Verhaltensgenerierung *effizient* erfolgt. Dies bedingt eine Minimalität der Prozesse und Repräsentationen. Insbesondere steht dies im Zusammenhang mit der Dichotomie von *Quantität* und *Qualität*. Qualitative Repräsentationen fassen einzelne Aspekte zu Einheiten zusammen, wohingegen quantitative Repräsentationen einzelne Entitäten entsprechend einer Metrik absolut spezifizieren. In diesem Spektrum, welches durch quantitative und qualitative Repräsentationen aufgebaut wird, sind solche Repräsentationen zu bestimmen, die eine effiziente Verhaltensfindung unterstützen. Eng verbunden mit der Effizienz der Prozesse und Repräsentationen ist die Forderung nach der *Effektivität* des Handelns, die danach fragt, welche Wirkung mit einer bestimmten, insbesondere einer kommunikativen Handlung erzeugt wird. Aus diesen beiden letzten Forderungen läßt sich der Begriff einer *optimalen Ressourcenadaptivität* umreißen. Ein Modell ist dann optimal ressourcenadaptiv, wenn die Verhaltensfindung maximal effizient erfolgt und dabei eine maximale Effektivität erreicht wird. Einerseits geht dadurch in diese Definition das prozeß- und repräsentationsorientierte Konzept der Effizienz ein. Andererseits wird der Nutzen eines Verhaltens durch das Konzept der Effektivität integriert. Besonders für kommunikatives Verhalten ist es notwendig, daß der Nutzen einer Beschreibung für einen Rezipienten zumindest ausreichend ist. Eine wenig detaillierte Beschreibung ist für einen Sprecher zwar effizienter zu bestimmen als eine detailliertere, jedoch kann deren Nutzen für den Rezipienten dennoch zu gering sein. In einem solchen Fall ist eine solche Beschreibung nicht optimal ressourcenadaptiv. Aus diesem Grund muß zwischen der Effizienz zur Bestimmung eines Verhaltens und dessen Effektivität ein Gleichgewicht hergestellt werden, welches i.a. domänenabhängig ist. In der Realität ist ein solches Optimum meist nicht erforderlich. Meist reichen auch suboptimale Approximationen aus, die sich effizient ermitteln lassen. Unter Berücksichtigung der vier Forderungen nach *Adaptivität*, *Flexibilität*, *Effizienz* und *Effektivität* wird in dieser Arbeit ein solches formales Modell entwickelt, welches als Verhalten inkrementelle Wegbeschreibungen erzeugt.

## 1.4 Gliederung

Die weiteren Kapitel sind entsprechend der verwendeten Methodologie *Empirie-Modellbildung-Implementation* aufgebaut.

In Kapitel 2 wird zuerst die Theorie behandelt, um einen Überblick über das gesamte Problemfeld der Raumkognition im allgemeinen und das von Wegbeschreibungen im speziellen aufzuzeigen und um die Terminologie einzuführen. Hierbei wird der aktuelle Forschungsstand hauptsächlich aus den Bereichen der Kognitiven Psychologie, der Ökologische Psychologie, der Anthro-Geographie, der Psycho- und Computerlinguistik und der Künstlichen Intelligenz systematisch im Hinblick auf die oben dargestellten Kernpunkte ausgewertet.

Daran anschließend werden in Kapitel 3 empirische Untersuchungen und deren Ergebnisse vorgestellt, die im Rahmen dieser Arbeit durchgeführt worden sind. Dieses Kapitel ist in zwei empirische Untersuchungen unterteilt: ein Feldexperiment und ein Laborexperiment.

In Kapitel 4 wird das formale Modell beschrieben, welches sich in den Rahmen der Theorie und

der Empirie einbettet. Es unterteilt sich in eine Vorstellung der Module und Funktionen und abschließend der Implementation. Das formale Modell, welches durch die Module und Funktionen beschrieben ist, wird als generischer Rahmen verstanden, welcher in der Implementation MOSES eine mögliche Realisierung gefunden hat.

Abschließend wird der Wert, den dieses Modell aus einer Ingenieursperspektive besitzt, in Kapitel 5 zusammen mit den offenen Fragen und einem Blick in die Zukunft diskutiert.

## Kapitel 2

# Theorien zur Raumkognition

Der den Menschen umgebende Raum ist seit altersher Gegenstand philosophischer, aber auch mathematischer Betrachtungen. Unser Verständnis für Raum hat sich seit Aristoteles über Euklid und Newton bis Einstein fortwährend geändert. Die Alltagsvorstellung von Raum ist nach wie vor stark von Newtons Konzept des homogenen, dreidimensionalen physikalischen Raums geprägt. Bezüglich dieses Raumes sind Kräfte von unterschiedlicher Reichweite und Stärke definiert. Dabei ist die Gravitationskraft diejenige, die das Leben des Menschen am stärksten bestimmt.

Solche formal eindeutig definierten Konzepte scheinen es jedoch nicht zu sein, die die Grundlage raumkognitiver Fähigkeiten bilden. Gesetze, wie beispielsweise das der Kommutativität von Distanzen werden vom Menschen nicht immer für die Lösung raumbezogener Probleme verwendet. In welcher Weise Menschen räumliche Sachverhalte repräsentieren, ist nach wie vor eine offene Frage. Einerseits sind geometrische Modellierungen weitaus genauer als das, was der Mensch verwendet. Andererseits reichen approximativ-qualitative Konzepte zur Modellierung menschlichen Verhaltens nicht aus. Es ist anzunehmen, daß der Mensch über verschiedene Repräsentationen verfügt, die nacheinander oder auch parallel zueinander automatisch oder auf Verlangen aufgebaut werden. Je nach Situation und Intentionen wird zwischen den einzelnen Repräsentationen des räumlichen Sachverhalts gewechselt.

Bei Langzeitrepräsentationen von räumlichem Wissen kommt es häufig zu *Verzerrungen* räumlicher Sachverhalte (vgl. [Tve81], [Tve92]). Ein solches Phänomen ist jedoch nur dann evident, wenn diese Repräsentationen mit physikalisch-mathematischen, meist geometrischen, Repräsentationen des Sachverhalts verglichen werden. Da der Mensch diesen Vergleich i.a. nicht vornimmt, tritt diese Diskrepanz für ihn nicht auf. Aus dieser Überlegung heraus können mentale Vorstellungen und physikalisch-mathematische Formalisierungen von räumlichen Sachverhalten als zwei unabhängige Sichtweisen auf gleiche räumliche Sachverhalte verstanden werden. Typischerweise zieht man sich bei einer Zerlegung des *externen* Raums auf eine physikalisch-mathematische Sichtweise zurück, wodurch letztendlich der Zugang zum externen Raum durch mathematische Konzepte gefiltert wird. Unter einer solchen Sichtweise lassen sich kognitive Repräsentationen und Funktionen formal untersuchen. Bei der Modellierung relevanter kognitiver Repräsentationen und Funktionen wird der physikalisch-mathematische Raum, ähnlich wie in der formalen Entscheidungstheorie, als optimale Vorgabe (base line) angesehen, auf die empirisch belegte Beschränkungen und Modifikationen angewandt werden. In einem solchen Paradigma ist es letztendlich notwendig, die Bedeutung und die Validität der physikalisch-mathematischen Repräsentationen zu überprüfen, um zu einer kohärenten Vorstellung über raumkognitive und allgemein kognitive Leistungen zu gelangen.

In welcher Weise räumliches Wissen abgeleitet, repräsentiert, evaluiert und genutzt wird, ist der Ansatzpunkt für eine Vielzahl von Forschungsrichtungen. Durch neurophysiologische Stu-

dien wird erforscht, welche Arten von Aktivitäten an welchen Lokationen des Kortex evoziert werden. Verschiedene Richtungen der Psychologie analysieren, wie höhere und niedere kognitive Systeme unter bestimmten mehr oder weniger eingeschränkten Bedingungen räumliches Wissen erwerben und nutzen. In der Anthro-Geographie steht hingegen im Vordergrund, wie der Mensch sich in seiner Umgebung orientiert und welche Eigenschaften der Umgebung in welcher Weise ausgenutzt werden. Kohärenzbildend werden Berechnungsmodelle der Künstlichen Intelligenz dazu verwendet, komplexe Zusammenhänge zwischen Funktionen und Repräsentationen im Zusammenspiel zu verstehen.

Allgemein akzeptierte Definitionen für raumkognitive Konzepte entstehen erst allmählich. Es lassen sich Aussagen über eine Vielzahl von Phänomenen machen, die jedoch nur in sehr seltenen Fällen auch Pendanten auf operationaler Ebene besitzen. Grundlage für die Untersuchung von Aspekten der Raumkognition ist eine Unterscheidung zwischen einem zu untersuchenden kognitiven System und einer dieses System umschließenden Umgebung. Je nach Schwerpunkt der Forschungsrichtung wird entweder das kognitive System oder aber die Umgebung stärker in den Vordergrund gerückt. Vielfach führt dies zu einer Vernachlässigung des kognitiven Systems oder der Umgebung. Zum Beispiel tritt in der Arbeit von Miller und Johnson-Laird, welche nach wie vor zum Standardwerk zur Integration von Wahrnehmung und Sprache gehört, der Begriff einer Umgebung in den Hintergrund. Stellvertretend sei hierzu ist das folgende Zitat:

“A theory of cognition should have two aspects. On the one hand, there should be a formal characterization of the functions or transformations performed by the cognitive system; on the other hand, there should be a descriptions of the psychophysiological mechanisms whereby such a system could be realized. Although there is considerable information about neural mechanisms underlying afferent processes, and a picture of some of the mechanisms underlying sensation and perception is beginning to emerge, the possibility of saying anything definitive about the conceptual mechanisms of the nervous system is sufficiently remote at the present time to justify omitting this aspect of a comprehensive theory from our considerations. At the level of sensation, however, enough work has been done, both in the formulation of laws relating sensory attributes to physical stimulation and in the analysis of processes in the sensory receptors and nerves, that we might expect to find the beginnings of a comprehensive theory.” (vgl. [MJL76, S. 12])

Typischerweise ist das Verhalten eines Akteurs zielgerichtet und orientiert sich am kurz- und langfristigen Nutzen von Verhaltensweisen. Intentionen steuern sowohl interne Funktionen, als auch Wahrnehmungsvorgänge und Verhaltensausführungen. Ebenso haben sie Einfluß auf die Evaluation raumkognitiver Funktionen und beeinflussen damit, *wann* welche räumlichen Repräsentationen *wie* verwendet werden. Drei Positionen kann ein Akteur bezüglich eines raumbezogenen Sachverhalts einnehmen. Erstens kann er als passiver Beobachter den Sachverhalt analysieren. Ebenso kann ein Betrachter innerhalb des Sachverhalts agieren und räumliches Wissen ableiten. Solch ein Wissen kann dazu verwendet werden, die räumliche Struktur des Sachverhalts zu erlernen und sprachlich zu beschreiben. In allen Fällen ist es notwendig, den Sachverhalt räumlich zu strukturieren, wozu das Konzept des räumlichen Referenzsystems eingeführt wird. In einem engen Zusammenhang stehen Referenzsysteme mit Prozessen der visuellen Objekterkennung. Mittels Referenzsystemen werden räumliche Sachverhalte und einzelne Objekte strukturiert (z.B. [Mar82], [Fel85]). Damit eng verbunden ist die Bestimmung von räumlichen Relationen, die *intraobjektiv* zwischen Objektteilen und *interobjektiv* zwischen Objekten etabliert werden. Dadurch lassen sich einerseits Objektteile zu einer gesamten Objektvorstellung (z.B. [Mar82], [Low87]) und andererseits Objekte miteinander zu Konfigurationen integrieren (z.B. [Kui78]). Bei der Speicherung von solcherart integriertem räumlichem Wissen unterscheiden

psychologische Theorien zwischen verschiedenen Speicher-„Medien“. Entsprechend der psychologischen Gedächtnistheorien wird Wissen im Kurzzeitspeicher (Kurzzeitgedächtnis) abgelegt, welches in einer Situation für die aktuelle Verhaltensfindung und -steuerung notwendig ist (vgl. [EZ94]). Teile dieses Wissens werden über Lern- und Behaltensfunktionen an den sogenannten Langzeitspeicher (Langzeitgedächtnis) weitergegeben. Durch Interaktion zwischen langsamem Langzeit- und schnellem Kurzzeitspeicher ist das kognitive System in der Lage, effizient und effektiv zu agieren. Der Kurzzeitspeicher wird als Vermittler zwischen Wahrnehmungs- und Verhaltensfunktionen sowie -repräsentationen verstanden.

Der Mensch verwendet räumliches Wissen, um sich in Umgebungen<sup>1</sup> zu orientieren und intentionsbasiert zu agieren. Eine solche phänomenologische Beschreibung gibt jedoch keinerlei Auskunft über die Verarbeitungsmechanismen und Repräsentationen.

Inhaltlich wird räumliches Wissen entsprechend seines entwicklungspsychologischen Auftretens in *Landmarken-* und *Routen-* sowie *Überblickswissen* aufgeteilt. Es wird allgemein davon ausgegangen, daß raumkognitive Fähigkeiten an sich nicht direkt, sondern nur über Mittlerfunktionen, insbesondere perzeptuelle Funktionen, mit einer externen Außenwelt in Kontakt treten. Über afferente Funktionen unterschiedlicher Modalität werden raumkognitiven Funktionen Informationseinheiten zugeleitet, die diese in räumliches Wissen transformieren. Eine der primären Quellen raumbezogener Information ist die visuelle Wahrnehmung. Über efferente Funktionen treten raumkognitive Funktionen mit solchen Modulen in Kontakt, die raumbezogene Information zur Verhaltenssteuerung verwenden. Vor allem Sprache und Motorik sind Empfänger solcher räumlicher Information.

Bei der Untersuchung raumkognitiver Fähigkeiten wird vielfach eine Unterscheidung nach dem Verhältnis eines externen Weltausschnitts zum Akteur vorgenommen. Ist der betrachtete Weltausschnitt größer als das, was der Agent von einer Lokation aus wahrnehmen kann, so spricht man von *large scale space* (vgl. [Kui78]). Im weiteren Verlauf wird dies als *Umgebung* bezeichnet. Der Weltausschnitt, den der Agent von einer Lokation aus vollständig perzipieren kann, heißt üblicherweise *small scale space* und wird im folgenden als *Situation* bezeichnet. Situationen sind demzufolge Ausschnitte aus Umgebungen.

Entsprechend dieser Unterscheidung in Umgebungen und Situationen differieren auch Methodologien zur Untersuchung raumkognitiver Leistungen. In laborbezogenen Untersuchungen werden vor allem situationsbezogene Erkenntnisse gewonnen. Dabei werden prinzipielle raumkognitive Leistungen, wie die Evaluation räumlicher Relationen zwischen abstrakten Entitäten, untersucht. Inwieweit Evidenzen aus solchen empirischen Studien in Bezug auf Realumgebungen ökologisch valide sind, ist meist eine offene Frage. Auf der anderen Seite des Spektrums werden in Umgebungen, die teilweise die Größe von Ländern haben können, solche Leistungen evaluiert, die bezüglich semantisch reichen Domänen verwendet werden. In solchen Untersuchungen ist es häufig schwierig, experimentalpsychologisch valide Hypothesen und Versuchsdurchführungen zu definieren. Andererseits kann dabei *natürliches*, d.h. alltäglich verwendetes Verhalten, evaluiert werden.

Eine Aufgabe raumkognitiver Leistungen ist es, aus Wahrnehmungsreizen extrahiertes, räumliches Wissen zu etablieren, miteinander zu verknüpfen, zu speichern und wieder zugänglich zu machen. Solcherart bestimmtes räumliches Wissen bezieht sich auf räumliche Lokationen, Verteilungen, Muster, Zusammenhänge, Konfigurationen und andere Entitäten, die einem bei der Bestimmung eines Pfades, der Aktivierung von Bewegungen, aber auch zur Beschreibung von räumlichen Sachverhalten dienen.

---

<sup>1</sup>Soweit es nicht anders vorgegeben wird, bezieht sich alles Folgende auf reale, d.h. nach physikalisch-mathematischen Gesetzen formalisierbare Umgebungen. Da u.a. aus ontogenetischen Gründen davon ausgegangen wird, daß der Umgang mit realem Raum die Basis für den Aufbau und die Verfeinerung von raumbezogenen Konzepten und Prozessen insgesamt ist, kann diese Annahme ohne Beschränkung der Allgemeinheit gemacht werden.

Um aus einer Situation Information auswählen zu können, verfügt der Mensch über Fokussierungsmechanismen. Ein solcher Mechanismus ist der der Aufmerksamkeit. In Bezug auf Aufmerksamkeit bezüglich visueller Wahrnehmung gibt es verschiedene Ansätze. Aus dem anfänglichen *Scheinwerfer-Modell* (vgl. [EE74]) entstanden die *Zoomlinsen-* (vgl. [ES86]) und die *Gradienten-Theorie* (vgl. [LB89]). Zu diesen Theorien sind zahlreiche Experimente zur raumbezogenen und zur objektbezogenen Aufmerksamkeit durchgeführt worden (vgl. [Dun84], [KH81], [KTG92], [KJ91]).<sup>2</sup> Die meisten Theorien zur visuellen Aufmerksamkeit nehmen diskrete Lokationen als Basis für Aufmerksamkeitsfunktionen an (z.B. [ES86], [LB89], [PC84], [TG80], [van92]). In diesem Kontext entwickelte Anne Treisman die lokationsbasierte *feature integration theory*, die auf einer merkmalsbasierten Indizierung von Lokationen beruht (vgl. [TG80], [Tre88], [Tre93] und Abbildung 2.1).

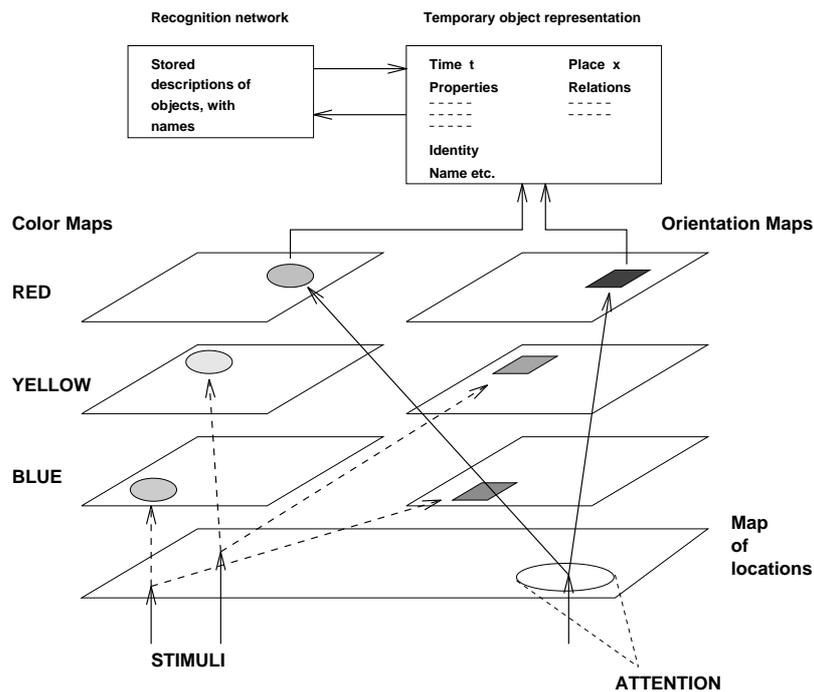


Abbildung 2.1: Feature Integration Theory (nach [Tre88])

In dieser Theorie nimmt Treisman an, daß verschiedene sensorische Merkmale, wie „Farbe“, „Orientierung“, „Größe“ und „Bewegungsrichtung“ in spezialisierten Modulen parallel und automatisch verarbeitet werden, während Objekte erst danach aus diesen Merkmalen mittels fokussierter Aufmerksamkeit identifiziert werden. Merkmale<sup>3</sup> werden automatisch und räumlich parallel verarbeitet, ohne daß dazu fokussierte Aufmerksamkeit notwendig ist (vgl. [Tre88, S. 203]). Eine solche Hypothese basiert auf anatomischen und physiologischen Befunden, in denen auf visuelle Kortexareale hingewiesen wird, die auf verschiedene Aufgaben spezialisiert sind (vgl. [Cow79], [MN87], [vM83], [Zek81]). Grundlage für die *feature integration theory* ist eine zweidimensionale, gerasterte *master map of locations* (siehe Abbildung 2.1). Von dieser zweidimensionalen Repräsentation werden für jeden Merkmalstyp durch Projektionen *Merkmalskarten* abgeleitet. Wenn Objekte gesucht werden, wird ein *Aufmerksamkeitsbereich* auf der *master map* etabliert und dieser Bereich an die Merkmalskarten weitergeleitet. In den einzelnen Merkmalskarten sind allein die Merkmalswerte, aber nicht deren Lokationen repräsentiert. Merkmale

<sup>2</sup>Für einen Überblick siehe [Log95].

<sup>3</sup>Treisman hält die Frage nach der Definition von Merkmalen für empirisch, die durch Konvergenzuntersuchungen beantwortbar ist (vgl. [Tre88, S. 230]).

im Aufmerksamkeitsbereich sind direkt zugreifbar. Je kleiner der Aufmerksamkeitsbereich ist, um so genauer ist die Information über die Lokationen und Werte der fokussierten Merkmale (vgl. [Tre88]). Falls ein Merkmal in einer Merkmalskarte einen Wert hat, der sich von den anderen Werten erheblich unterscheidet, zieht dieses Merkmal automatisch Aufmerksamkeit auf sich. Der *Auffälligkeitsgrad* eines Objektes ist im Unterschied der Merkmalswerte bezüglich einer Merkmalskarte begründet. Die *feature integration theory* basiert auf einem *bottom-up*-Prozeß. Treisman nimmt an, daß „bewußte“ Wahrnehmungen von temporären Objektrepräsentationen durch den Vergleich von verschiedenen Merkmalen und Verbindungen von Merkmalen mit im visuellen Langzeitgedächtnis gespeicherten Beschreibungen erfolgt (vgl. [Tre88, S. 204]). Dazu werden Strukturen aus der Integration von Merkmalskarten mit Langzeitrepräsentationen zueinander in Beziehung gesetzt. Nach Treisman bestimmt sich die Kapazitätsauslastung der Aufmerksamkeit aus der Anzahl der aktuell repräsentierten Objekte. Bei der Verarbeitung von Merkmalen und ihrer Verwendung zum Erkennen von Objekten gehen andere Modelle von integrierten “bottom-up” und “top-down” Funktionen aus (vgl. [LK79], [Zim80], [Rum77]).

Da in dieser Arbeit ein informationsverarbeitender Ansatz gewählt wird, ist es notwendig, eine Arbeitshypothese zu definieren, die es erlaubt, Merkmale zu klassifizieren und zu verarbeiten. Hierzu wird zwischen *rekursiven* und *linearen* Merkmalen unterschieden. Die Dekomposition der Projektionsebene in Lokationen bei der Analyse von visuellen Stimuli ist nur für die Verarbeitung *rekursiver* Merkmale sinnvoll. Ein Merkmalstyp wird in dieser Arbeit dann als *rekursiv* bezeichnet, wenn ein Träger eines Merkmals dieses Typs beliebig zerlegt werden kann, ohne daß einem dieser Teile der Zerlegung ein anderer Merkmalswert zugeordnet wird. Beispielsweise besitzt jeder Bruchteil einer homogen gefärbten Fläche dasselbe *Farbmerkmal* wie die Gesamtfläche. Die Merkmale *Höhe* und *Breite* besitzen diese Eigenschaft nicht und werden aus diesem Grund als *linear* bezeichnet. Zerschneidet man eine Fläche vertikal in zwei gleich große Hälften, so besitzt jeder Teil nur das Breitenmerkmal, welches halb so groß ist wie die Gesamtfläche. Addiert man die Breitenmerkmale beider Flächen, so erhält man den Wert des Merkmals der Gesamtfläche. Andere Beispiele für physikalische rekursive Merkmale homogener Objekte sind Wärme, Dichte, Helligkeit und Aggregatzustand. Weitere Beispiele physikalischer linearer Merkmale sind Masse und Strahlungsdichte. Rekursive Merkmale haben, im Gegensatz zu linearen Merkmalen, die Eigenschaft, nicht an die Einheit des Trägers gebunden zu sein. Aus diesem Grunde wird im weiteren Verlauf ein Verfahren zur Berechnung von visuellen Salienzen vorgestellt, welches für rekursive Merkmale auf einer Lokationsebene arbeitet. Für lineare Merkmale ist dies auf Grund der Objektbindung nicht möglich, so daß ein Verfahren auf Objektebene beschrieben wird. Visuelle Salienzen, die auf linearen und rekursiven Merkmalen basieren, werden letztendlich integriert und ermöglichen eine Gesamtaussage über die visuelle Salienz eines Objektes.

Bezogen auf raumkognitive Prozesse wird angenommen, daß Aufmerksamkeit notwendig ist, um räumliche Relationen zu berechnen (z.B. [EZ83], [Log95]). Danach werden keine räumlichen Relationen zwischen Objekten berechnet, die nicht im Aufmerksamkeitsbereich liegen. Aufmerksamkeit ist dazu notwendig, aus der prinzipiell unendlichen Menge von möglichen Relationen geeignete auszuwählen. Zwischen Entitäten im Aufmerksamkeitsbereich, werden räumliche Relationen solange nicht evaluiert, wie dies nicht intendiert ist (vgl. [Log95, S. 163]). Dies stellt einerseits starke Anforderungen an die Funktionen zur Berechnung räumlicher Relationen. Andererseits bewirken diese Beschränkungen eine Reduzierung der Komplexität.

Im alltäglichen Leben werden Beschreibungen räumlicher Sachverhalte vielfach verwendet. Dabei wird zwischen *realen*, perzipierbaren und *imaginierten* Räumen unterschieden.<sup>4</sup> Ausgehend von der ontogenetischen Entwicklung eines Menschen ist der reale, geographisch-physikalische Raum des Alltags als prototypisch anzusehen. Von diesem leiten Kinder raumbezogene Konzepte ab, die sie später auch auf abstrakte Domänen anwenden (vgl. [PI67]). Obschon in diesem Kontext

---

<sup>4</sup>Metaphorisch wird das Konzept des Raumes auch auf nicht-räumliche Domänen bezogen (vgl. [LJ80]).

eine Vielzahl von Einzelphänomenen untersucht worden sind, ist der konzeptuelle Rahmen der Raumkognition nach wie vor erst schemenhaft zu erkennen.

Natürlichsprachliche Beschreibungen dienen als effizientes Medium, mit dem ein Sprecher räumliches Wissen einem Rezipienten mitteilt. Dies bedingt, daß das Sprachsystem einen Zugang zur Ebene des räumlichen Wissens besitzt (siehe Abbildung 2.2). Konzeptuell lassen sich drei Ebenen unterscheiden. Auf der visuellen Ebene wird die visuelle Information, die der Betrachter aus der Situation selektiert und repräsentiert. Dies umfaßt u.a. die Prozesse zur Objekterkennung und der Objektverfolgung. Letztendlich wird die Situation auf der visuellen Ebene in Entitäten zerlegt, auf die andere kognitive Prozesse zugreifen können. Visuell erkannte Entitäten haben a priori keine räumliche Bedeutung für den Betrachter, solange er diese nicht in Beziehung zu sich und zu anderen Entitäten setzt. Bei der Betrachtung physikalischer Situationen sind zur Lösung von Aufgaben, bzw. zum Erreichen von Zielen, räumliche Lagebeziehungen von primärem Interesse. Dabei sind Lagebeziehungen zwischen dem Betrachter und einer Entität von solchen zwischen Entitäten zu unterscheiden. In Abbildung 2.2 sind Beziehungen zwischen Entitäten durch Kanten zwischen Knoten auf der Ebene des raumkognitiven Systems symbolisiert. Erst die räumliche Repräsentation ermöglicht es dem Betrachter eine Situation zu beschreiben. Entitäten und räumliche Relationen werden dazu auf Objektbeschreibungen und Beschreibungen von Lagebeziehungen abgebildet. Solche Abbildungen sind i.a. nicht eindeutig. Beispielsweise heißt dies, daß vielfach mehrere räumliche Relationen auf möglicherweise unterschiedlichen Abstraktionsebenen in ausreichender Weise gelten müssen, damit eine räumliche Präposition ausgewählt werden kann.

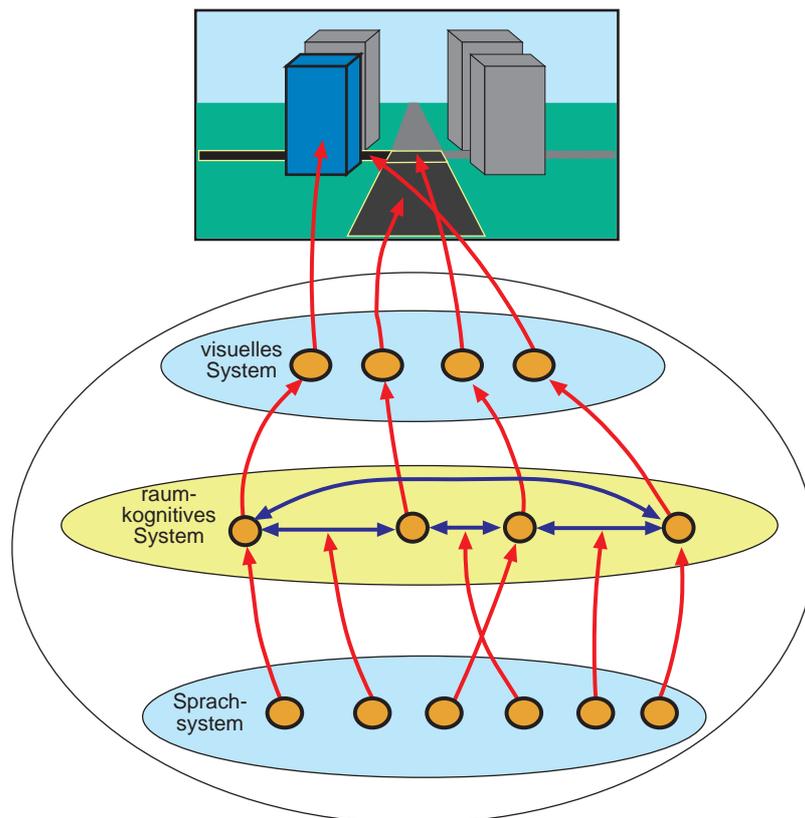


Abbildung 2.2: Ein Beispiel für die Verankerung sprachlicher Raumausdrücke in visuellen Daten

Dadurch sind sprachliche Äußerungen über räumliche Sachverhalte im räumlichen Wissen des Sprechers verankert. Dieses Wissen wiederum ist in der Information verankert, die dem Spre-

cher vor allem durch seine Perzeption zugänglich ist. In der Sprachforschung gibt es nur wenige Ansätze, die sich explizit mit einer solchen Verankerung sprachlicher Ausdrücke und Beschreibungen in perzipierten Daten auseinandersetzen. Auf Grund der formalen Klarheit sind denotationelle Semantiktheorien, wie Montague-Grammatiken (vgl. [Mon74]), von großem Einfluß. Dabei werden alleinig aus den sprachlichen Äußerungen Bedeutungseinheiten extrahiert und darüber Wahrheitswerte etabliert. Solche Theorien basieren auf Ansätzen der formalen Logik und ihrer Anwendung auf Sprache, so wie es von Frege als erstem vorgeschlagen worden ist (vgl. [Fre86b], [Fre86a]). Denotationelle Semantiktheorien sind nur bedingt für Raumbeschreibungen verwendbar, da sie nur indirekt über den Diskursbereich in perzipierten oder imaginierten Umgebungen verankert sind.

Ein anderer Vorschlag zur Konstruktion einer Semantiktheorie ist die *Situationssemantik* von Barwise und Perry (vgl. [BP83]). Sie geben einen allgemeinen Formalismus an, in dem sich relationale Beziehungen zwischen Entitäten in *einer* Situation und ihre Wirkung auf sprachliche Aussagen beschreiben lassen. Die Semantik einer Aussage ist dadurch nicht über Wahrheitsbedingungen oder Mengen möglicher Welten bestimmt, sondern über *Situationen* und *Beschränkungen* zwischen Situationen. Beschränkungen können formale aber auch naturgesetzliche Regelmäßigkeiten repräsentieren, wodurch Beschränkungen Bedeutung konstituieren (vgl. [Pin92]). Nomina werden durch Referenzrelationen in Objekten der Situation verankert.<sup>5</sup> Dadurch definiert sich beispielsweise die Bedeutung eines Bezeichners  $\beta$  im Satz  $\phi$  durch die Relation zwischen der Aussage, in der  $\beta$  vorkommt, und den Individuen, auf die  $\beta$  bezogen ist. Bisher konnte jedoch keine auch nur annähernd vollständige Theorie auf der Basis der Situationssemantik vorgestellt werden (vgl. [Pin92]).

Basierend auf einer geeigneten Semantiktheorie wird bei der Generierung einer Aussage konzeptuell unterschieden *was* und *wie* etwas gesagt werden soll (vgl. [Tho77]). Als *Konzeptualisierer* und *Formulierer* bezeichnet, ist diese Dichotomie aus psycholinguistischer Sicht untersucht und belegt worden (vgl. [KH87], [Lev89]). Eine ähnliche Unterscheidung wird in multimodalen Generierungssystemen verwendet (z.B. [FM89], [WAF<sup>+</sup>92], [Wah94]). Der Konzeptualisierer unterteilt sich nach Kempen und Hoenkamp in die Phase der *Makroplanung*, in welcher *Sprechakte* bestimmt werden, und in die Phase der *Mikroplanung*, in der *präverbale Bedeutungsstrukturen* festgelegt werden. Die *grammatikalische* und *phonologische Enkodierung* erfolgt im *Formulator* (vgl. [Lev89]). Dazu wird die präverbale Bedeutungsstruktur in *Oberflächenstrukturen* transformiert, welche dann in *Artikulationspläne* umgesetzt werden (vgl. [KH87]). In einer nachfolgenden Komponente, dem sogenannten *Artikulator*, werden Artikulationspläne ausgeführt (vgl. [Lev89]). Da ein Sprecher seine Aussagen selbst mitkontrollieren kann, nimmt Levelt eine *Sprachverstehenskomponente* an. In dieser analysiert der Sprecher seine *interne* Sprache und modifiziert gegebenenfalls die gesamte Äußerung (vgl. [Del80], [Lev83]). Eine präverbale Bedeutungsstruktur wird als eine propositional dargestellte semantische Repräsentation angesehen (vgl. [Lev89]). In den meisten Sprachgenerierungssystemen sind präverbale Bedeutungsstrukturen an die jeweilige Domäne angepaßt. Levelt schlägt eine hierarchische *Funktor/Argument*-Struktur vor, die aus dem Zusammenspiel verschiedener kognitiver Systeme entsteht (vgl. [Lev89, S. 78]). Dieser Ansatz basiert auf Jackendoffs *konzeptueller Semantik*, die eine Zwischenstufe zwischen allgemeiner Semantiktheorie und Sprachtheorie darstellt (vgl. [Jac83], [Jac87], [Jac90]). Levelt unterscheidet zehn Kategorien (vgl. Tabelle 2.1), aus denen sich präverbale Bedeutungsstrukturen zusammensetzen (ähnlich zu semantischen Netzen (vgl. [BS85]) und propositionalen Netzen (vgl. [And83])). Nur *EVENT* und *STATE* repräsentieren dabei vollständige Propositionen (vgl. [Lev89, S. 78]). Alle anderen sind elliptisch oder aber mit anderen Kategorien zu vollständigen zusammensetzbar. Beispielsweise können *THING* und *PLACE* die Dekomposition einer Kate-

---

<sup>5</sup>Die formale Repräsentation ist wie folgt strukturiert:  $d, c[\phi]e$  mit der Diskursituation  $d$ , der Referenzrelation  $c$ , Satz  $\phi$  und der Situation  $e$ , auf die sich die Äußerung bezieht.

Kategorie	Beschreibung
<i>PERSON</i>	Referenz auf eine Person
<i>THING</i>	Referenz auf ein Objekt
<i>EVENT</i>	Referenz auf ein Ereignis
<i>ACTION</i>	Referenz auf eine Aktion
<i>STATE</i>	Referenz auf eine Zustandsbeschreibung
<i>TIME</i>	Temporale Referenz
<i>PLACE</i>	Referenz auf eine Lokation
<i>DIRECTION</i>	Referenz auf eine Richtung
<i>ATTRIBUTE</i>	Referenz auf ein Attribut eines Objektes
<i>MANNER</i>	Referenz auf einen Modus einer Aktion

Tabelle 2.1: Kategorien der Bedeutungsstrukturen nach Levelt (vgl. [Lev89]).

gorie *STATE* sein.

Der Auftrag einer Beschreibung ist es, daß der Rezipient die *Intention* der Beschreibung versteht. Dies erfordert vom Sprecher, daß er die *Diskurssituation* ständig analysiert und relevante Veränderungen berücksichtigt. Die jeweils aktuelle Diskurssituation ist in ihrer Bedeutung für den Sprecher und den Adressaten vom *Diskurstyp*, wie z.B. Wegbeschreibungen, abhängig. Beschreibungen können vollkommen unverständlich sein, wenn zwischen dem Sprecher und dem Rezipienten keine ausreichende Übereinstimmung über den Diskurstyp vorhanden ist. Beispielsweise wird der Diskurs der vollständigen Wegbeschreibungen durch eine prototypische Frage wie folgt eingeleitet: „Entschuldigen Sie bitte. Können Sie mir vielleicht sagen, wie ich zum Stadion komme?“ Hierdurch wird dem Angesprochenen die Rolle des *Auskunftsgebenden* zugeordnet und zusätzlich festgelegt, daß beide sich nun in einem *Wegbeschreibungsdiskurs* befinden (vgl. [Kle79], [Wun78]).

Kognitive Beschränkungen des Sprechers, aber auch die vom Sprecher beim Rezipienten vermuteten, erfordern, daß der Inhalt einer Nachricht fokussiert wird (vgl. [GS86], [Her83]). Inhaltliche Sprünge sind dabei zu vermeiden oder so einzusetzen, daß sie vom Hörer nachvollziehbar sind. Wird ein wahrnehmbarer Sachverhalt beschrieben, so wird die Fokussierung teilweise durch diesen übernommen. Möchte der Sprecher bestimmte Aspekte eines Sachverhaltes hervorheben, so muß die Aufmerksamkeit des Hörers durch Beschreibungen<sup>6</sup> gelenkt werden.

Ist der Inhalt der intendierten Nachricht festgelegt, so wählt der Sprecher geeignete Informationseinheiten aus, von denen er erwartet, daß sie beim Rezipienten zum Verständnis der Aussage beitragen werden. Präferiert der Rezipient minimale Beschreibungen und leistet der Sprecher diesem Wunsch Folge, so wird der Sprecher eine weniger detaillierte Beschreibung generieren, als wenn der Sprecher annimmt, daß der Rezipient eine ausführliche Beschreibung bevorzugt. Als allgemeine Richtlinie können die Griceschen Konversationsmaximen verwendet werden, die vom Sprecher verlangen, daß er nur soviel beschreibt wie nötig und so wenig wie möglich (vgl. [Gri75]). In dynamischen Situationen sind auch die Ressourcen und hierbei besonders die *Zeit* zu berücksichtigen, die dem Sprecher für die Beschreibung zur Verfügung steht. Dadurch hat der Sprecher die Aufgabe, eine Beschreibung zu generieren und den gesamten Vorgang mit der Situation, in der er sich befindet, zu koordinieren.

Natürlichsprachliche Beschreibungen sind inhärent sequentiell, was erfordert, daß der Sprecher festlegen muß, in welcher Reihenfolge er die Nachricht mitteilen will. Dieses wird allgemein als

---

<sup>6</sup>Die Steuerung muß nicht in jedem Falle sprachlich sein. Im Alltag werden häufig Gesten verwendet, um die Aufmerksamkeit des Adressaten auf Bereiche im deiktischen Raum zu lenken.

*Linearisierungsproblem* bezeichnet (vgl. [Lev89, S. 138]). Verschiedenen Diskurstypen, wie beispielsweise Wegbeschreibungen, sind Prozesse assoziiert, über welche sich eine lineare Ordnung etablieren läßt. Das klassische Beispiel hierzu sind Wohnungs- (vgl. [LL75], [UE82]) und Wegbeschreibungen (vgl. [Kle79], [Wun78]), bei denen die Beschreiber das Konzept der *imaginären Wanderungen* verwendeten, um eine Beschreibung zu linearisieren. Ein ähnlich sequentialisierendes Konzept läßt sich in noch deutlicherer Form bei inkrementellen Wanderungen nachweisen (vgl. [Maa94]), in denen die Bewegung des Sprechers nicht nur die Reihenfolge, sondern auch die Zeitabstände zwischen Beschreibungen beeinflusst.

Die Mikroplanung besteht nach Levelt aus vier Elementen. Erstens ist jedem durch eine präverbale Bedeutungsstruktur referenzierten Objekt sein *Zugriffstatus* beigegeben, der angibt, ob das Objekt im aktuellen Fokus liegt. Zweitens kann der Sprecher die Aufmerksamkeit durch *Topikalisierung* der Beschreibung steuern. Das dritte Element betrifft die Wahl einer geeigneten *Perspektive*, bezüglich der die Beschreibung gegeben wird. In räumlichen Diskursituationen ist dies eng mit der Wahl eines geeigneten Referenzsystems verknüpft. Dies ist von prototypischen und kulturellen Faktoren abhängig. Wie in Kapitel 2.3.1 gezeigt wird, spielen für die adäquate Auswahl eines Referenzsystems eine Reihe allgemeiner kognitiver Prozesse eine Rolle.

Der *Formulator* besteht aus zwei Funktionen: die grammatikalische und die phonologische Kodierung. Die grammatikalische Kodierung transformiert präverbale Bedeutungsstrukturen in Oberflächenstrukturen. Einzelne Funktion/Argument-Strukturen aktivieren *lexikalische Einträge*. Diese wiederum aktivieren grammatikalische, morphologische und phonetische Generierungsfunktionen. Entsprechend der *lexikalistischen Hypothese* ist das Lexikon der Vermittler zwischen konzeptueller und grammatikalisch-phonologischer Kodierung (vgl. [Lev89, S. 181]). In der Phase der grammatikalischen Kodierung werden einzelne grammatikalische Strukturen, die von lexikalischen Einträgen abgeleitet worden sind, zu kohärenten und konsistenten Oberflächenstrukturen integriert. Diese werden durch die phonologische Kodierung in Artikulationspläne umgesetzt und dann anschließend zu einem geeigneten Zeitpunkt durch den *Artikulator* präsentiert.

Die bisher skizzierten Theorien zur visuellen Wahrnehmung, zum Gedächtnis und zur Sprachverarbeitung bilden einen groben Rahmen dessen, was ein Sprecher leistet, um raumsprachliche Beschreibungen zu generieren. Raumkognitive Funktionen basieren auf Leistungen, die durch diese ansatzweise formalisiert sind. Bisher gibt es keine Theorie der Raumkognition, die auch nur annähernd alle relevanten Elemente integriert. Im nachfolgenden Kapitel werden die Elemente diskutiert, die notwendigerweise in einer Theorie der Raumkognition enthalten sein müssen. Dabei wird ein besonderer Fokus auf den Aufbau räumlichen Wissens aus visuellen Daten und der Verwendung dieses Wissens für raumsprachliche Beschreibungen, insbesondere inkrementelle Wegbeschreibungen, gelegt.

Zum einen lassen sich Prozesse identifizieren, die das visuelle System betreffen. Die Schritte, welche notwendig sind, um ein Objekt zu erkennen, die Unterscheidung in *Hintergrund* und *Vordergrund*, der Aufbau von Referenzsystemen, anhand derer räumliche Lagebeziehungen zwischen Objekten und dem Wahrnehmenden und auch zwischen Objekten bestimmt werden können. Dies reflektiert eine physikalistische Betrachtungsweise. Nimmt man eine konzeptuelle Ebene zwischen Ein- und Ausgabemodulen an (siehe Abbildung 2.2), so stellt sich die Frage, wie die Repräsentationen auf beiden Ebenen einander zugeordnet werden. Miller und Johnson-Laird nehmen dazu eine pragmatische Sichtweise ein:

”One useful way of it is that perceptual judgments are the result of applying conceptual knowledge to sensory inputs.” (vgl. [MJL76, S. 60]).

Betrachtet man das raumkognitive System als abgegrenzten Teil des gesamten kognitiven Systems (z.B. [MJL76], [Bry92]), so gibt es *afferente* Funktionen, die raumbezogene Information

an das raumkognitive System liefern, und *effereute*, die sie weiter verwenden. Innerhalb des raumkognitiven Systems vereinen sich Fähigkeiten, wahrgenommene und imaginierte räumliche Information in räumliches Wissen zu transformieren und zu verarbeiten. Da es aus Effizienz- und Beschränkungsgründen nicht notwendig ist, permanent das gesamte räumliche Wissen aktiviert zu haben, ist es indiziert zugreifbar (z.B. [HJ85]).

Es lassen sich vier verschiedene Grundklassen räumliches Wissens unterscheiden:

1. *Objekte*: Entitäten, die durch perzeptuelle Fähigkeiten in einer Umgebung wahrgenommen worden sind, zuzüglich der räumlichen Ausprägung des Betrachters (ego) und des Rezipienten. Objekte sind in realen Umgebungen ein-, zwei- oder dreidimensional.
2. *Referenzsysteme*: Strukturierende Elemente räumlichen Wissens
3. *Räumliche Relationen*: Integrierende Elemente, die Objekte mittels Referenzsystemen qualitativ oder quantitativ zueinander in Beziehung setzen.
4. *Räumliche Konfigurationen*: lokale und globale räumliche Integrationen von Lokationen, bzw. Objekten mittels räumlichen Relationen bezüglich eines oder mehrerer, miteinander verbundener Referenzsysteme

Grundlegend für die Verwendung räumlichen Wissens ist das Konzept des „Objekts“. Aus der Sicht des Betrachters läßt sich die Lokation eines Objektes *qualitativ* oder *quantitativ* repräsentieren. Zwischen Objekten können räumliche Relationen unterschiedlichen Typs etabliert werden. Räumliche Relationen setzen i.a. das raumstrukturierende Element eines Referenzsystems als gegeben voraus (vgl. [MJL76]). Verschiedene Relationenklassen stellen unterschiedliche Anforderungen an Referenzsysteme. Empirisch lassen sich egozentrische und allozentrische Referenzsysteme nachweisen. Objekte werden durch räumliche Relationen auf der Grundlage eines Referenzsystems<sup>7</sup> miteinander in Beziehung gesetzt. Um die räumliche Struktur einer Situation zu erfassen und auch memorieren zu können, werden einzelne räumliche Relationen zu höheren räumlichen Strukturen integriert. Solche *räumlichen Konfigurationen* sind Grundlage für die Beschreibung einzelner Situationen und für den Aufbau räumlicher Langzeitrepräsentationen. Erst durch räumliche Konfigurationen wird eine Umgebung für den Betrachter gesamtheitlich kognitiv erfäßbar. Auf sprachlicher Ebene werden räumliche Relationen zumeist auf räumliche Präpositionen oder adverbiale Konstrukte abgebildet.

Grundlage für die Formalisierung raumkognitiven Wissens sind Konzepte der Euklidischen Geometrie, wie Achsen, Flächen, Richtungen, Distanzen und Lokationen im dreidimensionalen Raum. Dies beruht auf der Überlegung, daß die Euklidische Geometrie die wesentlichen Eigenschaften des Raumes, so wie ihn der Mensch erfährt, konzeptualisiert. Es läßt sich darüber spekulieren, ob die Euklidische Geometrie das Grundgerüst raumkognitiver Leistungen modelliert (vgl. [OB83, S. 250],[Jac87]). Aufgrund der intuitiven Plausibilität und der einfachen Formalisierbarkeit ist die Euklidische Geometrie die Grundlage für die meisten formalen Modelle im Kontext der Raumkognition.

Für den Aufbau räumlichen Wissens wird davon ausgegangen, daß eine Transformation von Wissensstrukturen perzeptueller und linguistischer Ebenen auf die der raumkognitiven Ebene stattfindet (z.B. [PI67], [Cas57], [Arn74], [MJL76], [OB83], [Bry92]). Die angenommenen Funktionen zwischen den Ebenen differieren in der Komplexität, wenn als Grundlage ein propositionales Format (vgl. [MJL76], [OB83]) oder einer Mischung aus einem propositionalen und einem depiktionalen Format (vgl. [Kos94]) angenommen wird.

<sup>7</sup>Topologische Relationen benötigen keine Verankerung in einem Referenzsystem.

Es wird angenommen, daß die Verankerung von sprachlichen Raumbeschreibungen in visuellen Daten (siehe Abbildung 2.2) durch Vermittlung eines oder mehrerer raumkognitiver Systeme erfolgt (z.B. [MJL76], [OB83]). In Anlehnung an Olson und Bialystok lassen sich auf den drei Ebenen verschiedene Prozesse klassifizieren (vgl. [OB83]):

1. *Perzeptuelle Ebene*: automatische und nur gering steuerbare Prozesse auf der niederen und höheren Ebene visueller Verarbeitung; nicht explizierbare Wissensstrukturen
2. *Raumkognitive Ebene*: artikulierbare und separierbare Einzelpropositionen, die intentional und gewollt gesteuert werden können; explizierbare Wissensstrukturen
3. *Sprachebene*: lexikalisierte räumliche Propositionen, die an den Diskurs angepaßt sind; explizierbare Wissensstrukturen

Perzeptuelle Prozesse sind nur in geringem Ausmaße steuerbar. Wissensstrukturen auf der perzeptuellen Ebene lassen sich nach Olson und Bialystok nicht explizieren (vgl. [OB83]). Perzeptuelle Information wird auf der raumkognitiven Ebene in räumliches Wissen transformiert. Sprachorientierte Funktionen haben auf dieses Wissen in ausreichendem Maße Zugriff.

Im weiteren wird die Unterscheidung zwischen Umgebungen und Situationen eingehend diskutiert, da sie für Wegbeschreibungen grundlegend sind. Die beiden Konzepte sind psychologischen Ursprungs und alleinig durch die Limitationen des Menschen gegeben. Anschließend wird detailliert vorgestellt, welche Raumkonzepte in Situationen relevant sind (vgl. Kapitel 2.2) und wie entsprechende Repräsentationen etabliert und verwendet werden (vgl. Kapitel 2.3).

## 2.1 Zusammenhang zwischen Umgebungen und Situationen

“The geographic world surrounding us is extremely complex. When we want to master a given problem in this world, we need to single out particular aspects of current interest from this multifaceted formation. So at any given time we are only interested in few objects, and concerning these objects again we are regarding only particular properties and/or relations. The capability of isolating the relevant aspects and relating them to one another, results in a unique intellectual efficiency. This efficiency, however, is necessary for successfully operating in the world”. (vgl. [FB94, S. 1]).

Der Begriff der *Umgebung* ist nicht eindeutig definiert, sondern stark von umgangssprachlichem Gebrauch geprägt. Von einem geographischen Standpunkt aus lassen sich verschiedene *physikalische Umgebungen* unterscheiden, wie beispielsweise Meeresumgebungen, Festlandumgebungen, Unterwasserumgebungen, atmosphärische Umgebungen, wetterbezogene Umgebungen und klimatische Umgebungen.<sup>8</sup> Ebenso lassen sich zwischen Umgebungen keine eindeutigen Grenzen ziehen, da sie meist ineinander übergehen.

Eine Umgebung ist durch *starke* und *schwache Dimensionen* bestimmt. Schwache Dimensionen sind solche, entlang welcher man sich mit nahezu konstantem Energieaufwand bewegen kann. Starke Dimensionen zeichnen sich dadurch aus, daß der Aufwand, sich auf dieser zu bewegen erheblich größer ist. Auffällig ist, daß wir in der alltäglichen Realumgebung keine Schwierigkeiten beim Wechsel von verschieden dimensionierten Umgebungen haben. Dies mag aber auch daher

---

<sup>8</sup>Solche natürlichen Umgebungen werden durch von Menschen geschaffene Umgebungen noch erweitert, was die Klasse dessen, was üblicherweise als Umgebung beschrieben wird, nahezu unendlich anwachsen läßt.

rühren, daß wir uns fast ausschließlich in zwei-dimensionalen, horizontalen Umgebungen fortbewegen (z.B. [Bol63], [Kru74]).<sup>9</sup> Obwohl die meisten Umgebungen des Alltags drei-dimensional sind, erfolgt die Fortbewegung maximal auf einer zweidimensionalen Ebene. Nur in seltenen Fällen spielt eine dritte Raumdimension eine Rolle. Auch wenn jemand einen Berg besteigt, so führt ihn sein Weg nur scheinbar in eine dritte Raumdimension, da die Ebene, auf der er sich fortbewegt, nur entsprechend der Steigung des Berges angepaßt ist. Vielmehr ist der Weg eines Bergsteigers nahezu durch nur *eine* Dimension bestimmt. Im alltäglichen Leben spielt aus diesem Grunde die dritte Dimension nur eine untergeordnete Rolle.<sup>10</sup> Aus diesen Überlegungen läßt sich schließen, daß eine Umgebung des alltäglichen Lebens fast ausschließlich durch zwei Raumdimensionen und zusätzlich einer Zeitdimension repräsentiert werden kann (vgl. [EM95]).<sup>11</sup>

Von Realumgebungen sind es städtische Umgebungen, mit denen der Mensch der westlichen Kultur am häufigsten konfrontiert wird (z.B. [Lyn60], [CS69], [App70]). Der Städteplaner Kevin Lynch klassifizierte städtische Umgebungen in fünf Kategorien (vgl. [Lyn60]). Die ersten drei Kategorien, *Landmarken*, *Wege* und *Verzweigungspunkte*, korrespondieren zu direkt perzipierbaren Gegenständen in einer Situation. Die zwei anderen Klassen, *Bezirke* und *Grenzen*, haben ihren Ursprung vor allem in physikalischen, politischen, sozialen, funktionalen und kulturellen Zusammenhängen.<sup>12</sup> Gleiches gilt für Bezirke. Sie sind im Laufe der Zeit gewachsen, haben sich aus unterschiedlichen Gründen verändert, und solange sie nicht normiert werden, sind sie nicht klar begrenzt. Bezirke sowie Grenzen sind für den einzelnen Menschen aus Erfahrung gewachsen und entstehen erst durch Gebrauch einer Umgebung. Dies korrespondiert mit der Unterscheidung in *Wahrnehmungsraum* und *Handlungsraum* (vgl. [Kru74, S. 57], für eine weitergehende Unterteilung siehe [Mon93]). Landmarken, Straßen und Kreuzungen können direkt wahrgenommen werden. Räumliches Wissen über Bezirke und Grenzen wird erst durch Interaktion, insbesondere durch Handlungen mit dieser Umgebung, schrittweise aufgebaut.

Im folgenden wird räumliches Wissen fokussiert, welches sich aus einem direkt perzipierten Weltausschnitt ableiten läßt.

## 2.2 Räumliches Wissen in Situationen

Die Anwendung von raumkognitiven Fähigkeiten erfolgt in Weltausschnitten unterschiedlicher Granularität. Ein Betrachter perzipiert einen räumlichen Sachverhalt von einem gegebenen Standpunkt immer vor dem Hintergrund seiner Ziele, Intentionen, Wünsche und Emotionen. Eine solche Perspektive auf einen bestimmten Ausschnitt einer Umgebung wird im weiteren als *Situation* bezeichnet (vergleichbar mit Kuipers "small scale space" [Kui78]). Ein Weltausschnitt, der im typischen Aktionsradius liegt, wird als *Umgebung* bezeichnet ("large scale space" (vgl. [Kui78])). Beispielsweise ist der Aktionsradius für einen Autofahrer um ein vielfaches größer als für einen Fußgänger. Um räumliches Wissen über eine Umgebung zu erlangen, muß der Akteur

---

<sup>9</sup>Dies ist möglicherweise ein Grund dafür, daß wir so fasziniert von Ballonfahrten und Tauchgängen sind, da sie uns ermöglichen, eine dritte Raumdimension zu erfahren. "Virtual Reality" hebt die Beschränkung in die dritte Raumdimension auf. Bei "Fahrten" durch 3D-Simulationen, in denen sich der Sitz auf dem man sich befindet, mitbewegt, geht dieses Schema der zwei Raumdimensionen verloren. Dies führt dazu, daß man Schwierigkeiten hat, sich an die vorgegebenen Wege zu gewöhnen. Erst nach einiger Zeit ist man in der Lage damit umzugehen, daß eine weitere schwache Raumdimension hinzugekommen ist, bzw. daß der Eindruck entsteht, eine starke sei zu einer schwachen transformiert worden.

<sup>10</sup>Steigt man eine Treppe hinauf, um zu einer anderen Etage zu gelangen, verwendet man zur Fortbewegung nur eine Raumdimension, nämlich die Höhendimension.

<sup>11</sup>Zweidimensionale Karten liefern i.a. eine hinreichend adäquate Repräsentation einer dreidimensionalen Umgebung. Eine wirkliche dritte Raumdimension kommt bei einem Tauchgang oder im Weltraum hinzu.

<sup>12</sup>Ein Beispiel, in dem eine Grenze physikalischer Natur ist, ist die Mauer zwischen dem zyprisch-griechischen und dem türkischen Teil von Nikosia. Anders ist dies mit fließenden Grenzen, wie beispielsweise dem Übergang vom Strand zum Meer. Dieser Übergang kann nicht exakt bestimmt werden.

räumliches Wissen einzelner Situationen miteinander integrieren.<sup>13</sup>

Eine Situation ist selten statisch, da sich zumindest entweder Objekte relativ zum Betrachter oder der Betrachter sich relativ zum Objekt bewegen. Durch Bewegungen verändert sich für den Betrachter die räumliche Struktur der Situation. Der Betrachter versucht, Situationsfolgen zu antizipieren, um bestimmte Vorhersagen über die Umgebung zu machen und adaptiert entsprechend sein Verhalten. Ändert sich eine Situation in einer unvorhergesehenen Weise, so muß der Betrachter sich in möglichst kurzer Zeit darauf einstellen, um so die Adäquatheit seines Verhaltens zu gewährleisten. In Situationen, über die der Betrachter Erfahrung besitzt, ist es ihm typischerweise ein Leichtes, sich auf Änderungen einzustellen (vgl. [Nei76]). In unbekanntem Situationen kann es dazu kommen, daß der Mensch nicht in der Lage ist, die Situationsfolge in geeigneter Weise zu antizipieren. Hierbei versucht er zumindest, solche Information herauszufiltern, die für die Verhaltensfindung von Bedeutung ist. Die Adaption erfolgt auf der Basis der Analyse externer Merkmale. Andererseits besitzt ein Betrachter *Intentionen*, die er in einer Situation in geeignete Verhaltensweisen umzusetzen sucht (vgl. [Bra81], [Bra84]). Unterschiedliche Intentionen haben zur Folge, daß der Betrachter auch unterschiedliche Objekte einer Situation fokussiert und identifiziert. So ändert sich u.a. die Menge der etablierten räumlichen Lagebeziehungen und darüber die mental etablierten räumlichen Vorstellungen.

Im weiteren werden allgemeine Konzepte der Raumkognition diskutiert. Fokussiert werden solche, die für den Aufbau räumlichen Wissens und dessen Verwendung in raumsprachlichen Beschreibungen von Situationen verwendet werden. Die grundlegenden Konzepte der Raumkognition sind die einer *Lokation*, eines *Objektes*, eines *Referenzsystems* und die der *räumlichen Relationen*. Eine Integration dieser Konzepte ermöglicht es, von einer Situation eine räumliche Vorstellung in Form räumlicher Konfigurationsbeschreibungen abzuleiten.

Im folgenden werden diese Elemente in einem größeren Rahmen eingeordnet, der räumliches Wissen feiner strukturiert.

## 2.3 Raumkonzepte

Die elementarste Entität des perzipierten Raums ist ein *Vorkommen* (vgl. [Gol91, S. 36]). Per Definition tritt ein Vorkommen in einem *Raum* auf. Jedes Vorkommen besitzt eine minimale Menge an Eigenschaften: *Identität*, *Lokation*, *Größenordnung* und *zeitliche Existenz* (vgl. [Gol91]). Über die Identität wird einem Vorkommen ein Name oder eine Markierung zugeordnet. Die Existenz eines Vorkommens in einer Situation wird durch eine Lokation indiziert. Eine Lokation ist dabei ebensowenig wie Identitäten ein festgelegter Begriff. Das Spektrum, bezüglich dessen die Lokation eines Vorkommens bestimmt werden kann, ist auf der einen Seite *quantitativ*, wie beispielsweise durch Angabe von Längen- und Breitengraden (z.B. im MERCATOR System, vgl. [Dav86]), und auf der anderen Seite des Spektrums *qualitativ*, durch Zuordnungen von nicht-metrischen Relationen (z.B. [RCC92], [FR93]). Eine grundlegende Eigenschaft einer Lokation ist, daß entsprechend dem „fundamentalen Gesetz der Geographie“, nie zwei Objekte gleichzeitig an einer Lokation sein können (vgl. [Tob76]). Lokationen sind weiterhin physikalische Merkmale zugeordnet. Die *Größenordnung* ist ein weiterer bestimmender Faktor eines Vorkommens. Sie hängt vor allem von kultur- und sprachabhängigen Einflüssen ab, die prototypische Erwartungshaltungen festlegen.<sup>14</sup> Die letzte Primäreigenschaft eines Vorkommens ist das der

<sup>13</sup>Die Unterscheidung in Situationen und Umgebungen ist vergleichbar mit Talmys Unterscheidung in *fiktive* und *faktive* Umgebungen (vgl. [Tal96]).

<sup>14</sup>Beispielsweise versteht man unter einem typischen Haus in Santa Barbara, Kalifornien, etwas grundsätzlich anderes als in Saarbrücken. Wohingegen man in Saarbrücken davon ausgeht, daß es sich um ein anderthalbstöckiges Gebäude aus Stein mit einem Spitzdach handelt, sind die Erwartungen in Santa Barbara eingeschossig mit Flachdach. Geschätztweise besteht zwischen beiden Prototypen ein Größenverhältnis von 2:1.

*zeitlichen Existenz*. Ein Vorkommen spannt dabei ein Spektrum auf, welches von *permanent* bis *kurzfristig* reicht. Vorkommen mit eindeutigen Identitäten, wohlspezifizierten Lokationen, verlässlichen und akzeptierten Größenordnungen sowie ausreichender Permanenz bezüglich des betreffenden Zeitbereichs scheinen die größte Kapazität in der Verankerung räumlicher Wissensstrukturen zu besitzen (vgl. Abbildung 2.3a und [Gol91, S. 38],[SW75], [PMS88]).

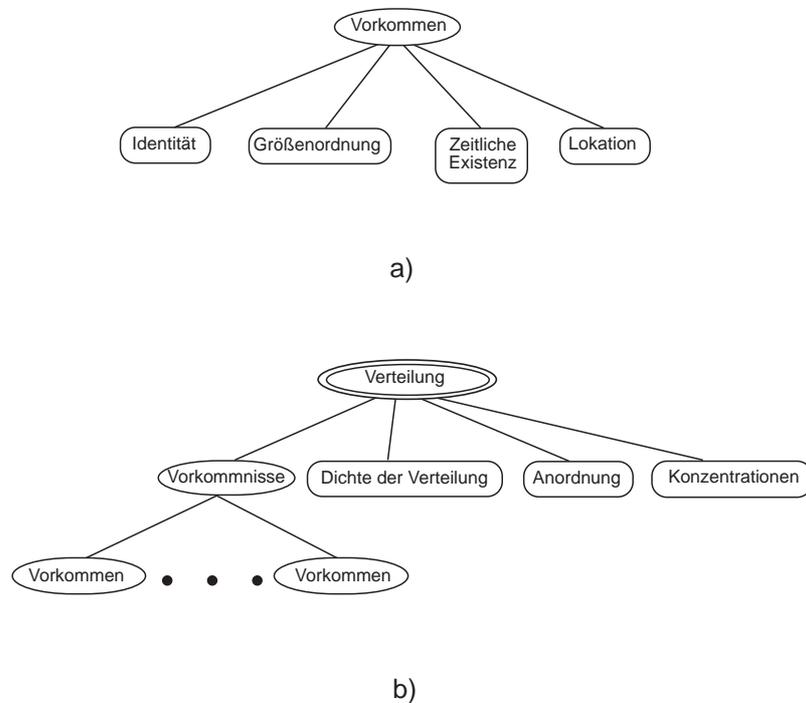


Abbildung 2.3: Hierarchie allgemeiner räumlicher Konzepte

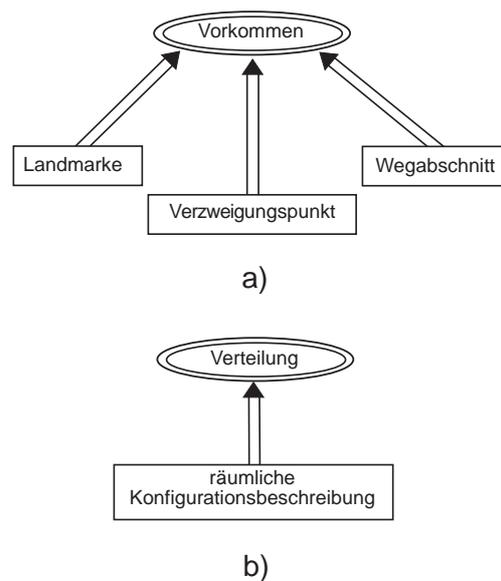


Abbildung 2.4: Konzepthierarchie einiger räumlicher Konzepte bezüglich städtischer Umgebungen

Weiterhin ordnen unterschiedliche Personengruppen Vorkommen unterschiedliche funktionale Eigenschaften zu, die sie relativ zu anderen Vorkommen als salient oder weniger salient erscheinen lassen. Auf der Basis von Faktoren lassen sich Vorkommen *kategorisieren* und *gruppieren*. Aufbauend auf dem Konzept der *Gruppierung* lassen sich Vorkommen zusammenfassen und in hierarchische Strukturen einer *Verteilung* integrieren (vgl. Abbildung 2.3b und [Gol91, S. 39]). Eine Verteilung kann wiederum selbst lokationsbezogene Eigenschaften, wie *Dichte der Verteilung*, *Anordnung* und *Konzentrationen* von Vorkommen, besitzen.

Vier grundlegende Raumkonzepte sind mit *Vorkommen* und *Verteilungen* assoziiert (vgl. [Gol91]):

1. *Distanz*: räumliche Nähe und räumliche Streuung
2. *Verbindung*: Kontakt, räumliche Sequenz und räumliche Ordnung
3. *Gebiet*: Inklusion, Exklusion und räumliche Menge
4. *Räumliche Stratifikation*: räumliche Hierarchien

Ein vom „fundamentalen Gesetz der Geographie“ abgeleitetes Axiom ist das der *räumlichen Separierbarkeit*, welches durch das Konzept der *Distanz* beschrieben wird. Damit verbunden sind die Konzepte der *räumlichen Nähe* bzw. der *räumlichen Streuung*. Ersteres ist ein Maß für die Nähe zwischen Vorkommen und Verteilungen und zweiteres ein Maß für die Distanz.<sup>15</sup>

Ein weiteres elementares räumliches Konzept ist das der *Verbindung*, welches sich auf räumliche Beziehungen zwischen Lokationen bezieht (vgl. [Gol91, S. 40]). Der Grad der Verbindung variiert von *vollständig* bis *fehlend*. Bezüglich dieser Gradierung werden die Konzepte *Kontakt*, *räumliche Sequenz* und *räumliche Ordnung* verwendet. Hierbei wird die inhärente Nähe zu Konzepten der Wissensrepräsentation und dabei besonders zu den Konzepten der *semantischen Nähe* und des “spreading activation” bei semantischen Netzen deutlich (vgl. [And83], [RM81]).

Das dritte Raumkonzept ist das des *Gebietes*. Hierüber werden diskrete Lokationen integriert. Elemente eines Gebietes sind räumlich untereinander enger miteinander verbunden als solche außerhalb dieses Gebietes. In Anlehnung an mengentheoretische Konzepte lassen sich bezüglich Gebieten Begriffe wie Inklusion, Exklusion und räumliche Menge definieren (z.B. [CGGT87], [HJ85]).

Ein viertes räumliches Konzept, welches auf sprachlicher Ebene von Bedeutung ist, ist das der *räumlichen Stratifikation*. Bezüglich allen Entitäten einer Umgebung werden räumliche Hierarchien verwendet, die eine Ordnung erzwingen.<sup>16</sup> Das Konzept der räumlichen Stratifikation ist auch ein organisatorisches Konzept für das Erlernen und Memorieren räumlichen Wissens (vgl. [HJ85]).

In verschiedenen entwicklungspsychologischen Theorien der Kognition, und insbesondere der Raumkognition wird zwischen verschiedenen Stufen der Aneignung räumlichen Wissens unterschieden (vgl. [SW75], [She62], [Gol78]). Primär ist danach die egozentrische Orientierung, die im Laufe der ontogenetischen Entwicklung zu absoluten Orientierungen erweitert wird. Entsprechend dieser Entwicklungsabfolge lassen sich auch unterschiedliche Wissensstrukturen unterscheiden. Der kognitionswissenschaftlichen Literatur folgend, die zwischen deklarativem und prozeduralem Wissen unterscheidet, wird das räumliche Wissen in *Landmarkenwissen* und *Routenwissen* unterteilt. Zusätzlich deuten die Ergebnisse aus zahlreichen empirischen Studien auf

<sup>15</sup>Ähnliche Konzepte werden durch Proximität, räumliche Variation, räumliche Ähnlichkeit bzw. Unähnlichkeit, Bereichsbildung, nächste Nachbarn und räumliche Heterogenität ausgedrückt (vgl. [Ans88]).

<sup>16</sup>Beispielsweise wird in Deutschland zwischen Katen, Ein-, Zwei- und Mehrfamilienhäusern, Villa und Palast unterschieden, die sich auf die Hierarchisierung von Behausungen beziehen.

eine dritte Form des räumlichen Wissens hin, welches als *Überblicks- oder Konfigurationswissen* bezeichnet wird (vgl. [SW75], [She62]).

Im weiteren wird zuerst darauf eingegangen, wie durch Etablierung verschiedener Referenzsysteme und räumlicher Relationen zwischen Objekten und Lokationen räumliches Wissen aufgebaut und sprachlich verwendet wird. Anschließend werden die drei Wissensarten Landmarken-, Routen- und Überblickswissen diskutiert, die sich von situationsbezogenem räumlichem Wissen ableiten lassen.

### 2.3.1 Räumliche Referenzsysteme

Um räumliche Sachverhalte zu repräsentieren, werden Vorkommen mit Hilfe von *räumlichen Referenzsystemen* strukturiert. Ein Referenzsystem kann sowohl von Eigenschaften eines Objektes oder einer Konfiguration von Objekten abgeleitet, als auch durch externe Normierungen vorgegeben werden. Zur Formalisierung von Referenzsystemen werden vor allem kartesische Definitionen verwendet (für einen Überblick siehe [Log95] und insbesondere [Her90], [MJL76]), wobei aber Modelle der visuellen Wahrnehmung auch auf die Bedeutung von *polaren* Referenzmodellen hinweisen (siehe [Pin84], findet auf raumkognitiver Seite Bestätigung durch [BS83], [Gal90]). Kartesische Koordinatensysteme spiegeln die Orthogonalität des menschlichen Körpers, entsprechend seinen Haupt- und Nebenachsen wieder. Gleichfalls spricht die sprachliche Einteilung des Raums entlang den kartesischen Dimensionen durch die Präpositionen „unten“, „oben“, „vorne“, „hinten“, „links“ und „rechts“ dafür.

Weiterhin sind im Kontext sprachlicher Beschreibungen räumlicher Sachverhalte die Verwendungsformen von Referenzsystemen vielfach betrachtet worden (z.B. im Kontext von „Deixis in Phantasma“ (vgl. [B34]), Versetzungsdeixis (vgl. [Sen85]) oder sekundäre Deixis (vgl. [Fil71], [Kle82], [Tal83]).

Basierend auf egozentrischen, intrinsischen und extrinsischen Referenzsystemen bestimmt der Betrachter seine Lokation bezüglich seiner Umgebung. Dies ist der Ausgangspunkt für den Aufbau raumkognitiver Repräsentationen (vgl. [HM73]). Entwicklungspsychologische Arbeiten zeigen, daß die Fähigkeit, Ordnungsrelationen bezüglich eines egozentrischen Referenzsystems zu etablieren, bei Kindern laut Piagets Studien im Alter von sechs Jahren zu entstehen beginnt. Erst im Alter von ungefähr acht Jahren erlernen Kinder die Fähigkeit, auch außerhalb des Körpers verankerte, intrinsische und extrinsische Referenzsysteme zu verwenden (vgl. [Pia26], repliziert durch [SB55]).

Eine elementare physikalische Eigenschaft, die für die Struktur eines Referenzsystems konstitutiv ist, ist das Konzept der *Gravitation G* (oben/unten-Achse, siehe Abbildung 2.5). Durch spezielle Funktionen des raumkognitiven Systems hat es der Mensch erlernt, Objekten, deren Hauptachsen nicht entlang der Gravitationsrichtung ausgerichtet sind, *allozentrische Referenzsysteme* zuzuordnen (vgl. [OB83]). Die zweite dominante Achse ist die *Links/Rechts-Achse*, welche orthogonal zur *Oben/Unten-Achse* parallel zum Gesichtsfeld bzw. den Projektionsebenen der Retinae verläuft (z.B. [HGH91]). Die dritte Achse wird durch die *Frontalrichtung* der aktuellen Aktion oder Perzeption festgelegt (Vorne/Hinten-Achse, siehe Abbildung 2.5). Eine solche Strukturierung des Raums definiert Regionen um das Zentrum eines Referenzsystems, weswegen es gebietskonstituierend ist (vgl. [Pri91]). In der horizontalen Ebene werden die Gebiete unterschieden, die *vor*, *hinter*, *links* und *rechts*, bzw. *nördlich*, *östlich*, *westlich* oder *südlich* vom Ankerpunkt des Referenzsystems liegen.

Es gibt Evidenzen, daß ein Akteur zu einem Zeitpunkt nicht nur über ein einziges, sondern über eine Hierarchie von Referenzsystemen verfügt, die er nach Bedarf etabliert und miteinander vergleicht (vgl. [MJL76], [OB83], [CRI93]). Referenzsysteme werden nach Levinson danach unterschieden, ob sie veränderlich (relativ) oder konstant (absolut) sind (vgl. [Lev96]). Absolute

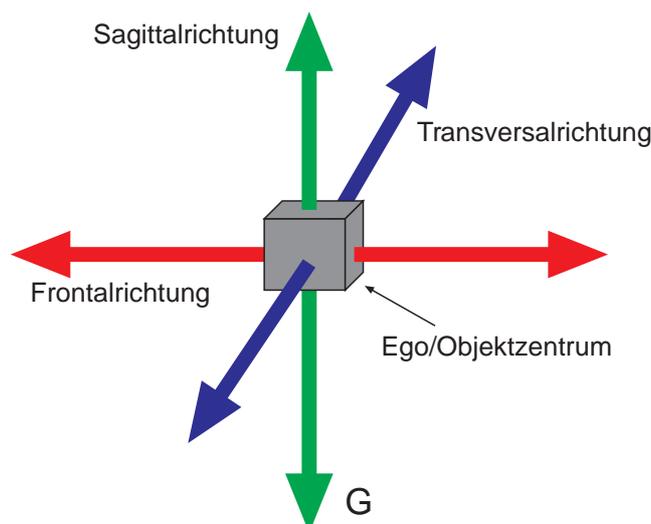


Abbildung 2.5: Modellierung eines Referenzsystems auf der Basis eines Kartesischen Koordinatensystems

Referenzsysteme, wie die Strukturierung des Raums durch Kardinalrichtungen, sind normativ und unabhängig von einzelnen Objekten und der Lokation des Betrachters. Relative Referenzsysteme hingegen werden bezüglich Entitäten oder Gruppen von Entitäten etabliert. Die Strukturierung des Raums, die durch ein relatives Referenzsystem erreicht wird, ändert sich, sobald sich die Entität in den Merkmalen ändert, die die Konstituierung des Referenzsystems bedingen. Die Entität, bezüglich der ein relatives Referenzsystem etabliert wird, kann entweder der Betrachter selbst oder eine Entität/Entitätsgruppe außerhalb des Betrachters (externe Entität) sein. Externe Entitäten wiederum lassen sich danach unterscheiden, ob sie Personen oder Objekte sind. Es gibt Evidenzen, daß es zumindest auf sprachlicher Ebene zu Unterschieden bei der Etablierung eines relativen Referenzsystems bezüglich eines Adressaten oder eines Objekts kommt (vgl. [Sch95b]).

Im weiteren wird nur die Klasse der relativen Referenzsysteme betrachtet, da ausschließlich diese im Korpus der inkrementellen Wegbeschreibungen nachgewiesen werden konnten.<sup>17</sup> Relative Referenzsysteme lassen sich in egozentrische und allozentrische unterteilen. Bei allozentrischen Referenzsystemen wird entsprechend ihrer sprachlichen Verwendungsweise zwischen intrinsischen und extrinsischen unterschieden:

1. *Egozentrische Referenzsysteme*: Sie sind im Sprecher verankert (z.B. [HBN87], [Ehr85], [Her86], [Gar89], [Lev84], [RS88], [BTF92]) und werden durch intrinsische Eigenschaften des Sprechers, bspw. in Abhängigkeit von der retinalen Projektionsfläche, der Ausrichtung des Kopfes oder Körpers determiniert (vgl. [MJL76], [CRI93]). Andere Bezeichnungen sind *deiktisch* oder *betrachterzentriert*.
2. *Intrinsisch definierte Referenzsysteme*: Sie sind allozentrische Referenzsysteme, die in einer perzipierten oder imaginierten Lokation, insbesondere einem Objekt oder anderen Personen, verankert sind (vgl. [Fil71], [Lev84], [Her86], [Gar89], [FL90], [Sch93]). Die Ausrichtung eines intrinsischen Referenzsystems wird durch Eigenschaften des Objektes determiniert.

<sup>17</sup>Da es ebenfalls zwischen Sprachen zu Unterschieden in der Verwendung von räumlichen Beschreibungen und von Referenzsystemen kommt, sei darauf hingewiesen, daß in dieser Arbeit nur deutschsprachige Wegbeschreibungen behandelt werden.

3. *Extrinsisch definierte Referenzsysteme*: Sie sind ebenfalls allozentrische Referenzsysteme und werden durch globale Eigenschaften einer Umgebung, wie einem Raum, einer Stadt oder einem Bild, oder durch den Betrachter determiniert (vgl. [OB83], [SH84], [CRI93]).

Egozentrische Referenzsysteme ( $for_e$ ) sind im *Egozentrum* eines Betrachters verankert (vgl. [HT66]).<sup>18</sup> Ein egozentrisches Referenzsystem unterteilt den Raum um den Betrachter in Gebiete bzw. im dreidimensionalen Raum in Volumina (siehe Abbildung 2.6 für einen 2D-Schnitt).

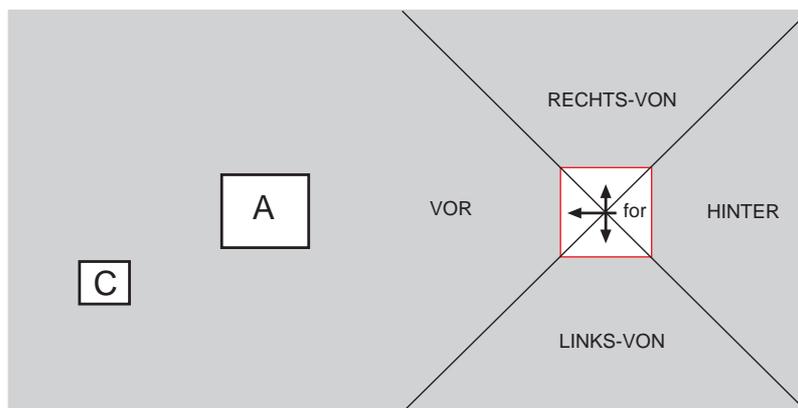


Abbildung 2.6: Durch ein egozentrisches Referenzsystem  $for_e$  konstituierte Gebiete in der horizontalen Ebene

Mittels intrinsischen Referenzsystemen werden Objekte zueinander in Beziehung gesetzt, ohne direkt Bezug auf das Egozentrum des Betrachters zu nehmen (z.B. [MJL76], [LJ93], [GM95], [Her90]). Es treten aber auch Fälle auf, in denen intrinsische Referenzsysteme in einer Lokation des imaginierten oder realen Raum verankert werden (z.B. [TT92], [Kle79], [Her86]). Im prototypischen Fall wird die Ausrichtung des Referenzsystems durch die Frontalachse festgelegt. Dies gilt insbesondere dann, wenn auf die Gravitationskraft Bezug genommen werden kann und die Transversalrichtung sich von der Frontalrichtung ableitet. Bei fehlender Gravitationskraft zeigt sich eine besondere Präferenzierung egozentrischer Referenzsysteme (vgl. [FL90]).

Verfügt das Referenzobjekt über *intrinsische* Eigenschaften, legt dies die Frontalrichtung des intrinsischen Referenzsystems a priori fest (siehe Abbildung 2.7). Es lassen sich vier wesentliche Fälle unterscheiden, die eine intrinsische Zuordnung ermöglichen (vgl. [MJL76], [Her90]):

1. Eine Seite des Objektes enthält *Perzeptionsfunktionen*
2. Die Seite, die zuerst *sichtbar* wird, wenn sich das Objekt bewegt (z.B. ein Auto)
3. Die *Standardposition* von Personen, wenn sie das Objekt oder die zugeordnete Objektklasse verwenden
4. *Situationsbedingte Eigenschaften*

Besitzt ein Objekt keine referenzierbare, intrinsische Eigenschaft, wird eine Frontseite von *außen* durch den Betrachter zugeordnet (vgl. [MJL76], [Her90], [CRI93]). Dies ist kulturell abhängig und im Deutschen durch die Seite bestimmt, die der Sprecher sieht. In manchen Kulturen wird der Seite die Eigenschaft *Front* zu sein zugesprochen, die auf der abgewandten Seite liegt (vgl. [Ped93]). Danach lassen sich Sprachen in solche unterscheiden, die ein Referenzobjekt nach einem *Spiegelbild-* bzw. nach einem *Tandemprinzip* ordnen (vgl. [Cla73], [Her86]).

<sup>18</sup>Der Ausdruck „egozentriert“ wurde von Roelofs eingeführt (vgl. [Roe59]).

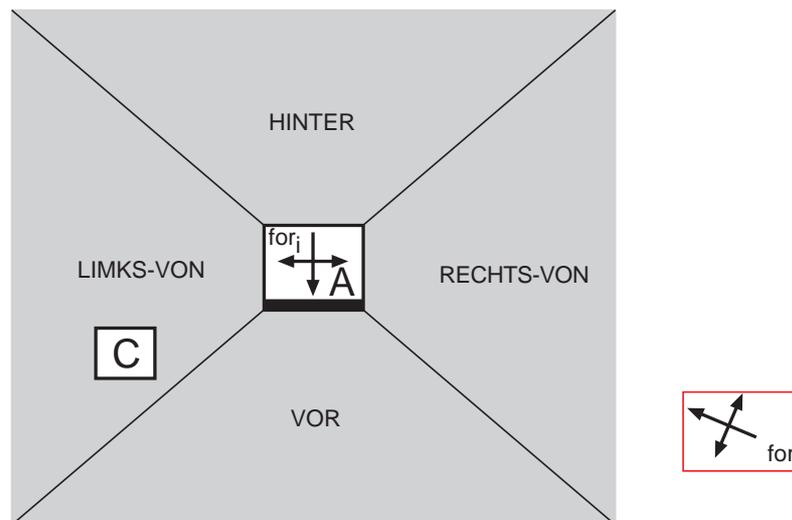


Abbildung 2.7: Durch ein intrinsisch determiniertes Referenzsystem  $for_i$  konstituierte Gebiete des Objektes  $A$  in der horizontalen Ebene. Der schwarze Balken in  $A$  gibt ein frontkonstituierendes Merkmal an.

Die horizontale Ausrichtung eines extrinsischen Referenzsystems wird vor allem durch den Kontext der Situation determiniert. Dies kann sowohl durch die räumliche Lagebeziehung zwischen dem Sprecher und dem Referenzobjekt gegeben sein (vgl. [Hil82], [MJL76], [Sai84]) als auch durch Eigenschaften der Umgebung, die den Sprecher und zu referenzierende Entitäten enthält (vgl. [CRI93]). In Abbildung 2.8 ist eine Situation illustriert, in der ein Betrachter eine räumliche Konfiguration aus einer schrägen Perspektive perzipiert und im Objekt  $A$  ein extrinsisch determiniertes Referenzsystem etabliert. Dabei ist die Front des Objektes  $A$  dadurch bestimmt, daß sie am nächsten zum Betrachter liegt (vgl. [Wun86], [Lev86]).

Auf der Basis dieser Referenzsysteme werden Umgebungen strukturiert. Je nachdem, welches Referenzsystem ein Sprecher einnimmt, generiert er Raumbeschreibungen durch *deiktischen*, *intrinsischen* oder *extrinsischen* Gebrauch von räumlichen Relationen (vgl. [MJL76], [RS88]<sup>19</sup> und Kapitel 2.5). Arbeiten über Referenzsysteme lassen sich in solche unterteilen, die eine deklarative Zuordnung von räumlichen Konfigurationen zu Referenzsystemen vornehmen (vgl. [MN78], [Fel87], [CRI93], [FL90]) und solche, die die Verwendung von Referenzsystemen als Grundlage raumsprachlicher Beschreibungen untersuchen (vgl. [Fil75], [Cla73], [MJL76], [Tal83], [Gar89], [FTC92], [LJ93], [HT95]). Es gibt Evidenzen, daß raumsprachliche Beschreibungen nicht nur jeweils auf einem Referenzsystem, sondern auf mehreren gleichzeitig beruhen. Es werden solche Beschreibungen präferiert, die in möglichst vielen Referenzsystemen verankert sind. Für die englischen Präpositionen “above” und “under” konnte dabei nachgewiesen werden, daß extrinsische Referenzsysteme gegenüber intrinsischen bei der Wahl einer vertikalen Orientierung präferiert werden (vgl. [CRI93]). Nach Carlson-Radvansky und Irwin hat die Verwendung eines egozentrischen Referenzsystems scheinbar nur geringe Auswirkungen auf die Wahl einer Beschreibung. Dies widerspricht in gewisser Weise den Ergebnissen, daß Umgebungen primär egozentrisch

<sup>19</sup>In diesen Arbeiten wird nicht deutlich, ob räumliche Relationen nur Kernbedeutungen räumlicher Präpositionen sind oder ob es sich um Repräsentationen von Lagebeziehungen zwischen Objekten handelt. Vor allem bei Retz-Schmidt wird zwischen verschiedenen Betrachtungsweisen gewechselt ([RS88]). Im ersten Falle ließe sich fragen, wovon räumliche Relationen abgeleitet werden. Dazu müßte eine weitere Klasse von sprachunabhängigen, räumlichen Relationen angenommen werden. Im zweiten Falle lassen sich räumliche Relationen als strukturelle Beziehungen zwischen Entitäten auffassen, so wie es visuelle Wahrnehmungstheorien postulieren (z.B. [Mar82], [Koe90], [MZ92]). Bisher ist diese Frage nach dem Status räumlicher Relationen weitgehend unbeantwortet.

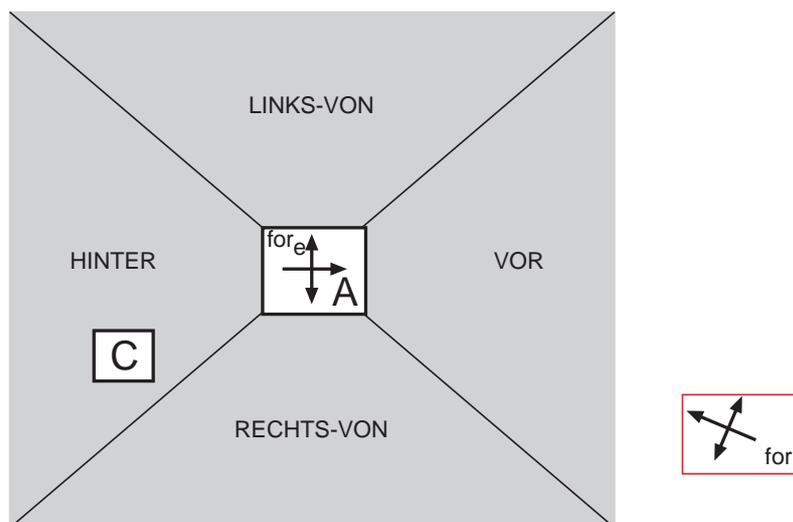


Abbildung 2.8: Gebiete, die durch ein extrinsisch determiniertes Referenzsystem  $for_e$  bezüglich Objekt A in der horizontalen Ebene konstituiert werden.

erfahren werden (vgl. [HBN87]). Aufgrund kognitiver Ökonomiekriterien ist alternativ die Interpretation möglich, daß egozentrische Referenzsysteme sowohl in die Bestimmung intrinsischer, als auch in extrinsische Referenzsysteme eingehen. Dadurch ließe sich erklären, daß egozentrische Referenzsysteme eine quasi „neutrale“ Wirkung besitzen.

Ein Problem bei der Untersuchung von Referenzsystemen besteht darin, daß sie i.a. mittels sprachlicher Beschreibungen evaluiert werden. Dabei sind Nebeneffekte zu identifizieren, die rein sprachlich-kultureller Natur sind. Entgegen den Evidenzen, daß es mit wachsendem Rotationswinkel schwieriger wird, sich in ein Objekt oder eine Person hineinzusetzen (vgl. [Her87], [HGH91]), wird in einem kommunikativen Kontext eine vis-à-vis Situation bevorzugt, in der die Gesprächspartner sich gegenüber sitzen (vgl. [Cla73], [Meh65], [Som69]). Offenbar kommt es bei der Verwendung von Referenzsystemen zur Bestimmung raumsprachlicher Beschreibungen zu Wechselwirkungen mit kommunikativen Präferenzen. Ein Sprecher möchte einem Rezipienten vielfach dadurch entgegenkommen, daß er dessen Perspektive, d.h., ein intrinsisches Referenzsystem annimmt (vgl. [Her87], [Sch93], [Sch95b]). Es ergibt sich beim Vergleich von Reaktionszeiten, daß in bestimmten Situationen eine intrinsische Beschreibung aus der Perspektive des Hörers länger dauert als aus der Perspektive eines Objektes (vgl. [Sch93], [Sch95b]). Da in diesen Untersuchungen keine explizite Unterscheidung zwischen einer Kommunikations- und einer Lokalisationsaufgabe gemacht wird, ist es nicht entscheidbar, ob dieser Effekt stärker kommunikativ-kulturellen Ursprungs ist, oder ob er auf der Basis eines neuen Referenzsystems erklärt werden muß.

In den meisten Studien zur Untersuchung von Referenzsystemen wird auf Effekte fokussiert, die auftreten, wenn ein Referenzsystem bereits ausgewählt worden ist (z.B. [MJL76], [Her86]). Es gibt bisher nur wenig Aufschlüsse über Mechanismen zur *Etablierung* von Referenzsystemen, zum *Wechsel* zwischen Referenzsystemen, sowie der *Verwaltung* von Referenzsystemen (z.B. [HGH91]). Aus linguistischer Sicht ist vielfach untersucht worden, welche Referenzsysteme konstituiert werden und wo diese relativ zur beschriebenen Situation verankert sind (z.B. [MJL76], [BTF92]). Es ergibt sich dabei eine Asymmetrie der Reaktionszeiten für räumliche Präpositionen der Frontalachse<sup>20</sup> für *intrinsische* Beschreibungen (vgl. [BTF92]). Diese Asymmetrie ist im

<sup>20</sup>Die Reaktionszeiten für eine Referenzierung durch “in front of” wird schneller gegeben als eine durch “behind”.

*extrinsischen* Fall nicht erkennbar. Beide Referenzsysteme sind in einer imaginären Lokation des Betrachters verankert, weswegen sie im erweiterten Sinne als egozentrisch angenommen werden können.

Zur Formalisierung von Referenzsystemen werden zuzüglich zu Eigenschaften der Referenzobjekte auch Information über die betrachterzentrierte, perspektivische Sicht des Sprechers verwendet, um Objekte in Gebiete zu zerlegen. (vgl. [Pri93, S. 178/9]). In einem anderen Ansatz wird entsprechend der Ausrichtung des egozentrischen Referenzsystems durch die geometrischen Eigenschaften eines Referenzobjektes ein umschließendes Rechteck definiert (vgl. [RS88], [HRA89]). Pribbenows Vorschlag ist es, das egozentrische Referenzsystem auf das Referenzobjekt zu induzieren und im egozentrischen Fall die Zuschreibungen der vorderen und hinteren Seite, entsprechend dem Spiegelprinzip (vgl. [RS88]) um  $180^\circ$  horizontal zu drehen (vgl. [Pri93, S. 178/9]). In beiden Fällen ist nicht ersichtlich, wie der Zusammenhang zwischen egozentrischen, intrinsischen und extrinsischen Referenzsystemen hergestellt wird. In diesen Arbeiten wird angenommen, daß die vertikale Ausrichtung a priori durch die Schwerkraft vorgegeben ist.

Bezüglich der Verwendung von Referenzsystemen gibt es Evidenzen, daß ein Sprecher mental durch Rotationen und Translationen allozentrische Referenzsysteme einnehmen bzw. aufbauen kann (vgl. [HBN87], [HGH91]). Es bietet sich zunächst an, Arbeiten zum Problem der mentalen Rotation heranzuziehen (z.B. [Cor82], [FS86], [How86]), in welchen die Fragestellung untersucht wird, wie ein Proband ein Objekt *A* mental derart rotiert, daß es deckungsgleich mit einem Objekt *B* ist. Im Gegensatz dazu wird bei Lokalisationsaufgaben eine Transformation des Blickpunktes des *Sprechers* und nicht eines Objektes mental untersucht.<sup>21</sup> In zwei Experimentalfolgen konnten Herrmann und Kollegen Effekte bei der Lokalisation unter Verwendung eines egozentrischen und eines in einem Objekt verankerten, allozentrischen Referenzsystems nachweisen, die nicht mit den Ergebnissen aus den Shepard-Metzler-Experimenten übereinstimmen (vgl. [HGH91]).

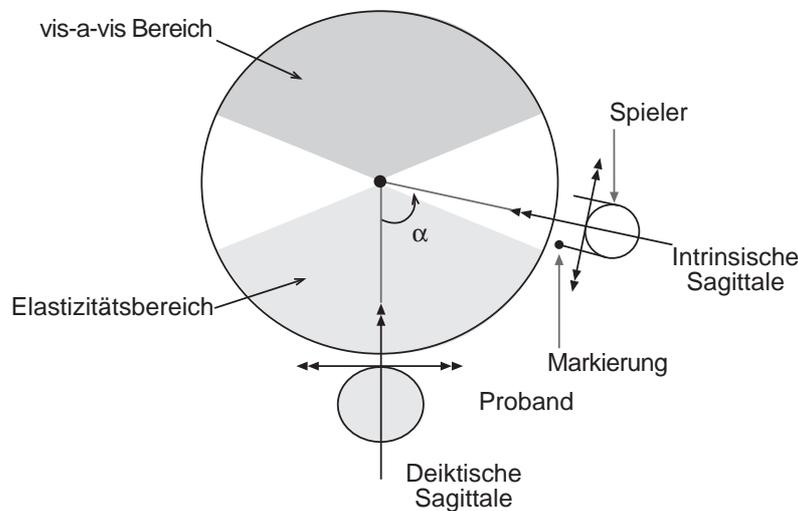


Abbildung 2.9: Versuchsanordnung und eingetragener Elastizitäts- und vis-à-vis-Bereich (vgl. [HGH91])

Intrinsische Lokalisation: Reaktionszeiten in Abhängigkeit vom Rotationswinkel  $\alpha$  (vgl. [HGH91]) [herrmann-intrin]

Es wurde getestet, wie schnell Probanden eine räumliche Lagebeziehung mittels „links“ oder „rechts“ verbalisieren konnten (siehe Abbildung 2.9). Der große Kreis symbolisiert einen Tisch,

<sup>21</sup>Herrmann et al. nennen dieses Phänomen das „Sichhineinversetzen“ (vgl. [HGH91]).

an dem Personen sitzen und Karten spielen. Zu jedem Zeitpunkt wird jeweils nur ein Spieler angezeigt. In einer Hand hält der Spieler seine Karten (Markierung). Die Ursprungsposition ist durch den Kreis unterhalb des Tisches angegeben. Es wurden zwei Versuchsreihen durchgeführt. In der ersten betrachtete der Proband den Tisch aus Position P heraus (*deiktische* Perspektive), bezüglich der er die Lage der Markierung anzugeben hatte ( $\alpha = 0^\circ$ ).<sup>22</sup> In der zweiten Versuchsreihe wurde der Proband aufgefordert, die Position des Spielers einzunehmen (*intrinsische* Perspektive).<sup>23</sup> Im deiktischen Fall wird das Referenzsystem<sup>24</sup> durch die egozentrische Frontalrichtung bezüglich Position P definiert; im zweiten Fall wird die intrinsische Frontalrichtung<sup>25</sup> bezüglich der Lokation des Spielers verwendet.

Das Ergebnis der deiktischen und intrinsischen Lokalisation zeigt, daß bei einem Winkel  $\alpha$  zwischen  $0^\circ$  und  $60^\circ$  die Reaktionszeiten nahezu konstant bleiben (siehe Abbildungen ??, 2.10). Ist  $\alpha$  größer als  $90^\circ$ , so steigt die Reaktionszeit im intrinsischen Fall stark an und erreicht ihr Maximum bei  $180^\circ$ . Im deiktischen Fall erreichen die Reaktionszeiten ein Maximum um  $120^\circ$  (vgl. Abbildung 2.10). Zwischen  $120^\circ$  und  $180^\circ$  kommt es wieder zu einem Rückgang auf ähnliche Reaktionszeiten, wie im Bereich zwischen  $0^\circ$  bis  $60^\circ$ . Dies unterstützt die Hypothese, daß sich die Probanden bei intrinsischer Lokalisierung bei kleinem  $\alpha$  (zwischen  $0^\circ$  und  $60^\circ$ ) ohne großen kognitiven Aufwand in den Spieler *hineinversetzen* konnten (Elastizitätsbereich). Schließen die

---

<sup>22</sup>In Abbildung 2.9 ist die Markierung bei einer deiktischen Beschreibung auf der *rechten* Seite.

<sup>23</sup>In Abbildung 2.9 ist die Markierung bei einer intrinsischen Beschreibung auf der *linken* Seite.

<sup>24</sup>In Abbildung 2.9 sind die Referenzsysteme in der horizontalen Ebene durch doppelköpfige Pfeile eingezeichnet. Das egozentrische Referenzsystem ist in der Lokation des Probanden und das intrinsische in der des Spielers verankert.

<sup>25</sup>In Richtung auf den Mittelpunkt des Tisches

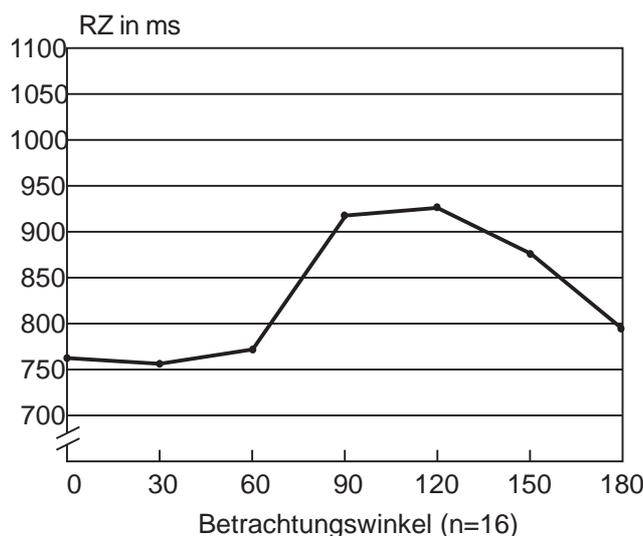


Abbildung 2.10: Deiktische Lokalisation: Reaktionszeiten in Abhängigkeit vom Rotationswinkel  $\alpha$  (vgl. [HGH91])

deiktische und die intrinsische Frontalrichtung einen größeren Winkel ein, so sind im intrinsischen Fall winkelproportional aufwendigere kognitive Leistungen vonnöten, um die Markierung zu lokalisieren. Im extrinsischen Fall wird um  $120^\circ$  ein Reaktionszeitenmaximum erreicht, woraus Herrmann et al. auf aufwendigere kognitive Leistungen für die Zuordnung des perzipierten Tokens zu einer der Markierungskategorien (*links* oder *rechts*) schließen. Im *vis-à-vis-Bereich* kommt es im deiktischen Fall zu einer Verringerung der Reaktionszeiten, da die Lokalisationsaufgabe ähnlich zu der im Elastizitätsbereich ist. Die Ergebnisse dieser beiden Versuche zur *Selbstrotation* stehen im Gegensatz zu der quasi-linearen Beziehung zwischen Rotationswinkel und Reaktionszeiten in den Untersuchungen mentaler Rotationen von Objekten. Aus dieser Experimentalserie läßt sich ableiten, daß die besten Leistungen zur Objektlokalisierung, sowohl im deiktischen, als auch im intrinsischen Falle, im Elastizitätsbereich von  $0^\circ$  bis  $60^\circ$  möglich sind. Diese Ergebnisse für intrinsische Beschreibungen sind von Schober repliziert worden (vgl. [Sch93], [Sch95b]).

Die meisten der genannten Arbeiten fokussieren die Verwendung von Referenzsystemen in statischen Situationen. Jedoch sind es meist dynamische Situationen, mit denen ein Sprecher im alltäglichen Leben konfrontiert wird. Durch Bewegungen werden Erwartungen für zukünftige Situationen vorgegeben. Eine Extrapolation der Trajektorie eines Objektes erlaubt die Antizipation einer Bewegung, so daß Situationen vorhergesagt werden können. Bewegt sich ein Sprecher real durch eine Umgebung, so nimmt er primär ein egozentrisches Referenzsystem als Grundlage seines Verständnisses von Raum an (vgl. [PM94], [Rie89]).

Zusammenfassend lassen sich mit absoluten und relativen Referenzsystemen zwei Klassen unterscheiden (vgl. [Lev96]). Absolute Referenzsysteme sind meist normativ in einer Umgebung festgelegt. Hingegen beziehen sich relative Referenzsysteme auf subjektive Erfahrungen und Eigenschaften des Betrachters. Im Kontext von Wegbeschreibungen im Deutschen werden ausschließlich relative Referenzsysteme verwendet, die sich weiter in egozentrische und allozentrische unterteilen lassen. Egozentrische Referenzsysteme bilden die Basis für deiktische Beschreibungen, wohingegen allozentrische Referenzsysteme für intrinsische und extrinsische Beschreibungen verwendet werden. Durch Festlegung eines Referenzsystems lassen sich räumliche Relationen zwischen Objekten der Situation etablieren.

### 2.3.2 Räumliche Relationen

Die Lokation eines Objektes wird für einen Betrachter erst dann verständlich, wenn sie durch *räumliche Relationen* zu einem perzipierten, memorierten oder bereits imaginierten Referenzobjekt in Bezug gesetzt worden ist. Räumliche Relationen lassen sich formal durch ein dreistelliges Prädikat explizieren ( $R(ego, lo, refo)$ ). Der Term *ego* repräsentiert die Lokation des Betrachters, wodurch ein direkter Bezug zwischen der Relation  $R$  und dem Betrachter hergestellt wird. In den meisten Arbeiten zu räumlichen Relationen wird dieser direkte Bezug zum Sprecher nur implizit betrachtet (z.B. [MJL76], [Her86]). Unter der Annahme, daß intrinsische und extrinsische Referenzsysteme von egozentrischen Referenzsystemen abgeleitet werden, ist es notwendig, die Lokation des Betrachters in der Formalisierung der räumlichen Relationen direkt zu berücksichtigen.<sup>26</sup> Der Betrachter ist der Ursprung für die Etablierung einer räumlichen Relation.<sup>27</sup> Der zweite Term *lo* steht für das *Lokalisationsobjekt*. Der dritte Term *refo* expliziert das *Referenzobjekt*, bezüglich dem das Lokalisationsobjekt in Beziehung gesetzt wird.<sup>28</sup> Eine solche Figur-Grund-Eigenschaft von räumlichen Relationen zeigt eine gewisse Nähe zu Eigenschaften bei der Separierung von Objekten im Kontext der visuellen Wahrnehmung.

Nach Pullar und Egenhofer [PE88] lassen sich räumliche Relationen in vier Klassen unterteilen:

1. *Topologische Relationen*: Lagebeziehungen zwischen zwei Objekten, die *Disjunktheit*, *Kontakt*, *Überlappung* oder *Inklusion* explizieren; topologische Relationen benötigen kein Referenzsystem (für Formalisierungen vgl. [Ege91], [EAT92], [Coh87], [RCC92], [Rup96])
2. *Deiktische Relationen*: repräsentieren Lagebeziehungen zwischen dem Sprecher und Objekten (vgl. [MJL76], [RS88])
3. *Ordnungsrelationen*: Lagebeziehungen, die Objekte anhand einer Ordnung sequenzialisieren
4. *Metrische Relationen*: explizieren den metrischen Aspekt einer Lagebeziehung, wobei die verwendete Metrik vom Sprecher und Hörer in gleicher Weise verstanden sein muß (z.B. [McN92])

Eine andere Unterscheidung ist es, räumliche Relationen in *quantitative* und *qualitative räumliche Relationen* einzuteilen (vgl. [FR93]). Metrische Relationen sind rein quantitativ, wohingegen die ersten drei Klassen räumlicher Relationen qualitativ sind. Quantitative räumliche Relationen sind *lokationskonstituierend*, wohingegen qualitative räumliche Relationen *gebietskonstituierend* sind. Qualitative Relationen geben keine genaue Lokation, sondern ein Gebiet an, welches prototypisch dieser Relation in einer gegebenen Situation zugeordnet wird. Ein Beispiel für eine quantitative Relation ist eine Relation  $R_1$ , die repräsentiert, daß zwei Lokationen  $x$  und  $y$  sich genau in einem Abstand von einem Meter befinden. Ein Beispiel für eine qualitative Relation ist eine solche, die auf sprachlicher Ebene durch die Präposition „nahe“ ausgedrückt wird.

<sup>26</sup>Ein weiterer Vorteil dieses Ansatzes ist es, daß räumliche Relationen in ihrer Stelligkeit nicht variieren, sondern homogen behandelt werden können.

<sup>27</sup>Bei Sekundärdeixen wird die Lokation des Sprechers *ego* durch eine Referenz zu einer Lokation in dieser lokalen Umgebung ersetzt (siehe [Kle79]). In der Beschreibung „Schaut man von Berkeley aus nach San Francisco, so liegt Palo Alto in einer Relation  $L$  von San Francisco.“ nimmt der Sprecher eine imaginäre Position ein, die in diesem Falle ganz Berkeley umfaßt. Es gilt somit die Relation  $L(\text{loc}(\text{Berkeley}), \text{loc}(\text{Palo Alto}), \text{loc}(\text{San Francisco}))$ . Zwischen der Lokation, die durch  $\text{loc}(\text{Berkeley})$  repräsentiert wird, und dem Sprecher gibt es dadurch einen kausalen Zusammenhang, daß sich der Sprecher mental in die Lokation von Berkeley versetzt. Es gilt somit zwischen der Lokation des Sprechers und Berkeley zusätzlich eine Transitionsrelation  $S(\text{ego}, \text{loc}(\text{Berkeley}))$ . Dadurch gilt die Relation  $L$  dann, wenn der Sprecher seine Lokation *ego* imaginär nach Berkeley verlagert. In diesem Falle wird  $L$  durch eine *links-von* Lagebeziehung repräsentiert.

<sup>28</sup>In der Figur-Hintergrund-Terminologie wird *lo* als *Figur* und *refo* als *Hintergrund* bezeichnet (vgl. [Tal83]).

Für jede Klasse von räumlichen Relationen werden unterschiedliche Anforderungen an das verwendete Referenzsystem gestellt (siehe Tabelle 2.2). Topologische Relationen benötigen kein Referenzsystem, da die Ausrichtung und Lokation vom Betrachter unabhängig ist (vgl. [MJL76], [Ege91]), wohingegen deiktische Relationen auf qualitative Eigenschaften eines egozentrischen Referenzsystems zurückgreifen (vgl. [MJL76], [RS88]). Ordnungsrelationen und metrische Relationen können bezüglich allen Referenzsystemen etabliert werden. Dabei nehmen Ordnungsrelationen eine lokale oder globale Ordnungseigenschaft bezüglich der Dimensionen eines Referenzsystems an. Metrische Relationen setzen voraus, daß durch das Referenzsystem *Metriken* bezüglich den Dimensionen definiert sind. Eine Eigenschaft einer solchen Metrik ist es, Distanzen zwischen Lokationen entsprechend einer Funktion vorzugeben, was für einen Betrachter ohne Meßinstrumente oder ausreichende Erfahrung i.a. schwierig herzustellen ist. Eine Metrik ordnet jedem Paar zweier Lokationen  $x$  und  $y$  eine Distanz  $d(x, y)$  zu, welche drei Bedingungen erfüllt (vgl. [McN92]):

1. Positivität:  $d(x, x) = 0$  und  $d(x, y) > 0$ , wenn  $x \neq y$
2. Symmetrie:  $d(x, y) = d(y, x)$
3. Dreiecksungleichung:  $d(x, z) \leq d(x, y) + d(y, z)$
4. Additivität der Segmente:  $\forall$  Lokationen  $x, z$ :  $x$  und  $z$  sind durch ein Segment verbunden:  
 $\forall y$  auf dem Pfad zwischen  $x$  und  $z$ :  $d(x, z) = d(x, y) + d(y, z)$ <sup>29</sup>

Metrische Relationen werden im räumlichen Wissen typischerweise verzerrt repräsentiert. Es kommt u.a. zu Asymmetrien der Distanzen zwischen Objekten (vgl. [Ste69], [Cad79], [SBS80]) und zur Verletzung der Additivitätsbedingung.

	Egozentrisches Referenzsystem	Allozentrisches Referenzsystem	ohne Referenzsystem
Topologische Relationen			•
Deiktische Relationen	•		
Ordnungsrelationen	•	•	
Metrische Relationen	•	•	

Tabelle 2.2: Beziehung zwischen räumlichen Relationenklassen zu Referenzsystemklassen

Deiktische Relationen werden wesentlich durch egozentrische Referenzsysteme konstituiert. Es werden durch diese Relationen in einer dreidimensionalen Umgebung sechs Gebiete identifiziert: VOR, HINTER, ÜBER, UNTER, RECHTS-VON, LINKS-VON (siehe Abbildung 2.6). Assoziiert mit diesen Gebieten ist eine beschränkte Menge von deiktischen Relationen: #in-front-of#, #behind#, #above#, #under#, #right-of#, #left-of#. <sup>30</sup> Dadurch läßt sich angeben, welche Bedingung erfüllt sein muß, damit eine deiktische Relation  $DR$  zwischen einer imaginierten oder realen Lokation des Sprechers  $S$  und einem Objekt  $O$  erfüllt werden kann (vgl. Definition DEIKTISCHE RELATIONEN).

#### Definition 1: Deiktische Relationen

Eine deiktische Relation  $DR(\text{loc}(S), \text{loc}(O), \text{loc}(S))$  ist erfüllbar, gdw.  $O$  im zu  $DR$  assoziierten Gebiet bezüglich eines egozentrischen Referenzsystems liegt.

Ordnungsrelationen lassen sich in ähnlicher Weise über Gebiete bezüglich eines Referenzsystems etablieren (siehe Abbildung 2.7, 2.8 und Definition ORDNUNGSRELATIONEN).

Bei metrischen Relationen reduziert sich die Erfüllbarkeit auf Lokationen (vgl. Definition METRISCHE RELATIONEN). Da sich Gebiete i.a. nicht eindeutig voneinander abgrenzen lassen, werden in manchen Arbeiten sogenannte *Anwendbarkeitsgrade* eingeführt, die angeben, inwieweit die Lokation eines Objektes in einer *prototypischen Region* des Gebietes liegt (vgl. [Her86],

<sup>29</sup>Diese Bedingung wird eingeführt, um triviale Metriken, wie z.B.  $d(x, x) = 0$  und  $d(x, y) = 1$ , auszuschließen. Eine Distanzmessung, die alle vier Bedingungen erfüllt, wird als "Metrik mit additiven Segmenten" bezeichnet.

<sup>30</sup>Die Restriktion der räumlichen Relationen erfolgt entsprechend den Prinzipien der minimalen Repräsentation und dem geringsten Zugeständnis. Eine solche Einschränkung wird auch durchgängig in der psychologischen Literatur vorgenommen (z.B. [GW95], [Log95], [HB92]).

---



---

Definition 2: Ordnungsrelationen

Eine Ordnungsrelation  $OR(\text{loc}(S), \text{loc}(O1), \text{loc}(O2))$  ist erfüllbar, gdw.  $O1$  im zu  $OR$  assoziierten Gebiet bezüglich eines in  $O2$  verankerten Referenzsystems liegt.

---



---



---



---

Definition 3: Metrische Relationen

Eine metrische Relation  $MR(\text{loc}(S), \text{loc}(O1), \text{loc}(O2))$  ist erfüllbar, gdw.  $O1$  auf der zu  $MR$  assoziierten Lokation bezüglich der Metrik  $M$  eines in  $O2$  verankerten Referenzsystems liegt.

---



---

[ABHR87], [Gap94], [Her96]). Ein solcher Ansatz erweist sich besonders bei der Umsetzung auf eine sprachliche Ebene als vorteilhaft (z.B. [ABHR87], [Her96]).

Mit Objekten, räumlichen Referenzsystemen und räumlichen Relationen sind die Konzepte eingeführt worden, die für den Aufbau räumlichen Wissens notwendig sind. Im folgenden wird diskutiert, wie im Kontext von Raumbeschreibungen geeignete Referenzsysteme ausgewählt (Abschnitt 2.3.3) und verschiedene Arten räumlichen Wissens aufgebaut werden (2.3.4, 2.3.5, 2.3.6).

### 2.3.3 Wahl eines Referenzsystems im Kontext von Raumbeschreibungen

Wissen über Lokationen von Objekten in einer Umgebung ist eine notwendige Voraussetzung für adäquates sprachliches Verhalten. Um auf eine Lokation Bezug nehmen zu können, wird diese Lokation zu der bereits etablierten räumlichen Struktur in Beziehung gesetzt.

Drei Bedingungen sind für eine erfolgreiche Kommunikation räumlicher Relationen notwendig (vgl. [Kle94]).

1. Der Sprecher und der Rezipient teilen sich in einem ausreichend hohen Maße einen kognitiv-rekonstruierten Raum
2. Der Sprecher und der Hörer beherrschen die lexikalischen Bedeutungen raumbezogener Ausdrücke
3. Die Kernbedeutung wird mit allen Sorten von Kontextinformation passend unterstützt (z.B. durch Positionen und Blickrichtungen)

Unter der Annahme, daß der zweite Punkt eine Grundvoraussetzung ist, wird jetzt der erste Punkt eingehender betrachtet. Inwieweit Kontextinformation verwendet werden kann, wird in Kapitel 2.5.2 diskutiert.

Eine Voraussetzung für die Verankerung räumlichen Wissens ist die Auswahl geeigneter Referenzsysteme. Um einen räumlichen Sachverhalt kommunizieren zu können, muß der Rezipient Referenzsysteme annehmen, welche denen des Sprechers hinreichend ähnlich sind. Die Wahl eines geeigneten Referenzsystems ist für den Nutzen einer Beschreibung entscheidend. In der sprachorientierten Literatur finden sich u.a. sechs Faktoren, die die Wahl eines Referenzsystems beeinflussen:

1. Verbale Explikation (vgl. [Sai84], [MJL76])
2. Intrinsische Präferenz (vgl. [MJL76], [Abk82], [Cox81], [Lev82], [Wun82])
3. Individuelle Unterschiede (vgl. [Lev82])
4. Konstante intrinsische oder deiktische Beschreibung in einem Diskurs (vgl. [LL75], [vC93], [Ehr85])
5. Statischer oder dynamischer Charakter einer Situation (vgl. [Wun81])
6. Integrierte Betrachtung von situationsbedingten Faktoren und intrinsischen Eigenschaften des Referenzobjektes (vgl. [GM95])

Im ersten Fall wird die Art des Referenzsystems dem Hörer explizit mitgeteilt, so daß die Wahl willkürlich ist. Vielfach untersucht worden sind im Hinblick auf die Wahl eines Referenzsystems, in welcher Weise intrinsische Eigenschaften dies beeinflussen. Obschon sie einen starken Einfluß haben, kann es dennoch zu Ambiguitäten kommen (vgl. [Her90]), die u.a. kulturabhängig sind (vgl. [Car93], [Kle94]). Die Kategorisierung in Personengruppen, die deiktische, und solche, die intrinsische und extrinsische Beschreibungen bevorzugen, erscheint fragwürdig (vgl. [GM95]). Weiterhin lassen sich Kulturabhängigkeiten nachweisen (z.B. in [Hav92], [Lev92], [Ped93]). In gleicher Weise ist eine diskursbedingte, konsistente Verwendung einer Klasse von Referenzsystemen nur in Spezialfällen, wie Wohnungsbeschreibungen, belegbar (vgl. [LL75]). Die Zuordnung von Referenzsystemen entsprechend dem Charakter von Situationen läßt sich nur vage aufrecht erhalten. Vielversprechender ist der letzte Faktor. Grabowski und Weiß konnten zeigen, daß in einer institutionalisierten Situation in unserer Kultur allozentrische Referenzsysteme bevorzugt werden (vgl. [GW95, S. 12]). In informellen Situationen werden hingegen meist egozentrische Referenzsysteme verwendet. Allgemein wird in der deutschen Sprache in den meisten Fällen eine deiktische Beschreibung gewählt, um ein Referenzobjekt zu indizieren, welches keine intrinsischen Eigenschaften besitzt (vgl. [GM95]). Besitzt das Referenzobjekt intrinsische Eigenschaften, so beeinflussen der *situative Kontext* und die *intrinsischen Merkmale* des Objektes gemeinsam, ob ein deiktisches, intrinsisches oder extrinsisches Referenzsystem verwendet wird (vgl. [GM95]). Andere Arbeiten gehen davon aus, daß die Etablierung eines intrinsischen oder extrinsischen Referenzsystems mit erheblichen kognitivem Aufwand verbunden ist und daß aus diesem Grund egozentrische Referenzsysteme präferiert werden (z.B. [BNH86], [Ehr85], [HBN87], [Hil82], [Lev82], [Lev84, Wun81], [CRI93]). Interessant ist, daß einfache raumbezogene Beschreibungen vom Rezipienten zu einem großen Anteil nicht korrekt verstanden werden.<sup>31</sup>

Auf der Basis von geeigneten Referenzsystemen verbindet ein Betrachter Objekte einer Umgebung durch Etablierung von räumlichen Relationen zu räumlichen Wissenstrukturen. Es werden dabei drei räumliche Wissenstrukturen unterschieden, die sich auf das Wissen über Landmarken, Routen und ganzen Konfigurationen beziehen. In den nächsten Abschnitten werden diese nacheinander diskutiert.

### 2.3.4 Landmarkenwissen

In jeder Umgebung gibt es Objekte, die in räumlichen Beschreibungen als markant oder auffällig beschrieben werden. Durch Verwendung solcher Objekte strukturiert ein Betrachter seine Umgebung. Solche markanten Objekte werden als *Landmarken* bezeichnet. Lynch war der erste, der

<sup>31</sup>In empirischen Untersuchungen konnte von Grabowski und Miller gezeigt werden, daß bezüglich einer Einparksituationsbeschreibung nur 63% aller raumbezogenen Beschreibungen zu einer erfolgreichen Kommunikation führten, in der sich der Hörer und der Sprecher verstehen (vgl. [GM95]). Für die deutsche Sprache zeigt sich, daß Hörer und Sprecher über das Gebiet zwischen beiden und dem Referenzobjekt erheblich besser kommunizieren als über den Raum hinter dem Referenzobjekt (vgl. [GM95]).

eine Landmarke als eine raumstrukturierende Einheit auffaßte (vgl. [Lyn60]). Hiernach ist eine Landmarke ein Referenzpunkt, zu dem andere Objekte einer Umgebung in Beziehung gesetzt werden, wodurch sie der Wegfindung in komplexen Umgebungen dienen (vgl. [App70], [BW76], [CS69], [Lyn60], [ASR78]). Eine Voraussetzung ist die *räumliche Prominenz* eines Objektes, die dadurch entsteht, daß das Objekt von mehreren Seiten sichtbar ist oder sich von anderen Objekten der Situation bezüglich seiner Merkmale abhebt (vgl. [Lyn60]). Subjektive Einstellungen des Betrachters modifizieren die Prominenz eines Objektes. Liegt ein Objekt beispielsweise in der Nähe einer Kreuzung, an welcher der Betrachter abbiegen möchte, steigert dies die Prominenz des Objektes (vgl. [Lyn60]). In perzipierten Umgebungen ist die Prominenz wesentlich durch visuelle Eigenschaften eines Objektes determiniert (vgl. [PM88]). Die Zuordnung von referentiellen Eigenschaften erlaubt es, andere Lokationen relational mit Landmarken in Beziehung zu setzen (vgl. [SBS80]).

In der psychologischen Literatur wird zwischen Landmarken unterschieden, die in einer Situation direkt perzipiert werden (z.B. [Gib79]) und solchen, die strukturierende Bedeutung in räumlichen Langzeitrepräsentationen besitzen (vgl. [Pre82], [Pre87], [PM88]). Im prototypischen Fall wird eine Landmarke als ein stationäres Objekt klassifiziert, welches eine Lokation im Raum bestimmt (z.B. [App69], [CS69], [SW75], [PMS88]). Andererseits können auch bewegliche Objekte wie die Sonne<sup>32</sup> als Landmarken fungieren. Entwicklungspsychologisch lassen sich zwei Sichtweisen auf Landmarken identifizieren (vgl. [HM73], [SW75]). Landmarken dienen der Orientierung, der Wegfindung und dem Verfolgen eines Weges (vgl. [Acr88]).

Das Piagetsche Konzept einer *topologischen Repräsentation* ist in diesem Kontext so interpretiert worden, daß Kinder Lokationen nur unter Verwendung von Landmarken, und dabei insbesondere räumlich naher Landmarken, referenzieren können (vgl. [APO75], [HS78]). Da nur dann eine Aussage über die räumliche Nähe zwischen zwei Objekten gemacht werden kann, wenn vorher der Abstand zwischen diesen beiden bestimmt worden ist, tritt das Paradoxon auf, daß räumliche Nähe, bzw. Nachbarschaft, inhärent mit Distanzen und darüber mit Metriken verbunden sind. Diese offene Frage kann dadurch umgangen werden, daß angenommen wird, daß Kinder, im Widerspruch zu Piaget, über ein metrisches Konzept der Distanz (vgl. [LGS81]), bzw. ein nicht-metrisches Konzept der Nachbarschaft oder eine Fähigkeit der Ordnungsbildung verfügen (vgl. [New88]). Die Bedeutung von Landmarken für die räumliche Orientierungsfähigkeit wird durch neurophysiologische Studien unterstützt. Durch Läsionen des *ventralen System*<sup>33</sup> bedingte Probleme, Objekte zu erkennen, haben auch Auswirkungen auf das Erkennen von bekannten Landmarken (vgl. [LWF85]) und dadurch auf die Fähigkeit, Wege zu finden (vgl. [Sho92]).

Aus Gründen der Ökonomie erscheint es wahrscheinlich, daß nicht alle potentiellen Relationen zwischen Landmarken automatisch etabliert werden (vgl. [Sad88]). Ein effizienterer Ansatz ist es, Relationen nur zwischen ausgewählten Referenzpunkten und Landmarken zu bestimmen, wodurch räumliches Wissen hierarchisiert wird (z.B. [SC78], [Tve81], [HJ85]). Dies ist konsistent mit dem Vorschlag, einer Landmarke ein minimales Referenzsystem zuzuordnen (vgl. [PYR79]), wodurch sich Lokationen und Landmarken zu Konfigurationen zusammenfassen lassen (z.B. [PMS88]). Durch Zuordnung eines Referenzsystems werden Landmarken zur Orientierung und Lokalisierung verwendet (vgl. [PM88], [Nad88]). Eine solche Sicht basiert auf der Annahme, daß raumkognitive Fähigkeiten inhärent relational sind (vgl. [PM88]). Landmarken bewirken typischerweise eine Verzerrung des memorierten räumlichen Wissens. So werden Distanzen zwi-

---

<sup>32</sup>In dem Sinne, daß die meisten, möglicherweise nahezu alle, Menschen, entgegen des physikalischen Wissens, die Sonne immer noch als sich um die Erde drehend erfahren.

<sup>33</sup>Es gibt die Vermutung, daß das visuelle System aus zwei getrennten Systemen besteht, welche in Areal V1 des visuellen Kortex verankert sind. Das ventrale System endet im Temporallappen und verarbeitet in wesentlichen Teilen Information, die sich auf die Eigenschaften von visuell wahrgenommenen Objekten bezieht. Das dorsale System hingegen endet neurophysiologisch im Paritallappen und verarbeitet hauptsächlich räumliche Information (vgl. [MUM83]). Plakativ werden diese beiden als die "What"- und "Where"-Systeme bezeichnet.

schen Landmarken und Nicht-Landmarken als größer angenommen als umgekehrt (vgl. [Cad79], [HM82], [SBS80]).

Für die raumkognitive Integration von Landmarken wird ein Kriterium benötigt, nach dem Relationen ausgewählt werden. Ein geeignetes Kriterium ist das der *räumlichen Nähe*. Räumliche Nähe ist eng mit dem Konzept der Distanz verbunden (vgl. [Pri93], [MJL76], [Gla94]). Über Distanzen wird bestimmt, ob zwei Lokationen als räumlich nahe oder getrennt betrachtet werden. Dies korrespondiert zum „ersten Gesetz der Geographie“, wonach alles mit allem verbunden ist, jedoch nahe Entitäten mehr miteinander verbunden sind als entferntere (vgl. [Tob70]). Das Konzept der räumlichen Nähe modelliert diese Proportionalität der Abschwächung der räumlichen Verbundenheit mit wachsender Distanz. Bezogen auf raumkognitive Fähigkeiten besitzt das Konzept der räumlichen Nähe weniger eine metrische als eine qualitativ-ordnende Bedeutung (vgl. [Gol93]). In diesem Zusammenhang ist eine Reihe von Modellen zur Formalisierung qualitativer Distanzen entwickelt worden (z.B. [Fra91], [FO91], [Zim93]). Das Grundkonzept ist dabei die Verwendung von unscharfen, gewichteten Gebieten anstelle von punktförmigen Lokationen. Zusammenfassend kann gesagt werden, daß Landmarkenwissen statisches, räumliches Wissen über Umgebungen, insbesondere von Objekten, repräsentiert. Über Landmarken werden Objekte zu anderen Objekten der räumlichen Vorstellung des Betrachters in Beziehung gesetzt. Neben Landmarkenwissen eignet sich ein Betrachter dynamisches, räumliches Wissen, sogenanntes Routenwissen, an. Routenwissen wird von Aktionen abgeleitet, die in einer Umgebung ausgeführt werden. Die Grundlage für Aktionen sind Intentionen bzw. Ziele, bezüglich derer sie ausgeführt werden.

### 2.3.5 Routenwissen

Bei zielbezogenen Handlungen werden einzelne Lokationen entlang eines Pfades von einem Start- zu einem Zielort erreicht. In den meisten Fällen sind solche Pfade, wie z.B. durch Straßen, vorgegeben. Das Wissen, das man benötigt, um zielgerichtet von einem Ort zu einem anderen zu gelangen, wird als *Routenwissen* bezeichnet (vgl. [SW75]). Entsprechend entwicklungspsychologischer Modelle enthält Routenwissen keine Information über relative Beziehungen zwischen Lokationen, die nicht einem Routenabschnitt zugeordnet sind. Routenwissen entsteht dadurch, daß mit Landmarken und Pfadelementen prozedurale Information assoziiert wird (vgl. [SW75, S. 24-29]). Dazu wird bezüglich eines *Verzweigungspunktes* vermerkt, welche Aktion dort typischerweise ausgeführt wird. Der Begriff des Verzweigungspunktes ist ähnlich dem Konzept eines *Entscheidungspunktes* (vgl. [Edw54]) oder *Aktionspunktes* (vgl. [Bre92], [Ker96]) aus dem Bereich der Entscheidungstheorie. Im weiteren wird unter einem Verzweigungspunkt eine Lokation verstanden, an dem entschieden werden muß, in welche Richtung ein Pfad weiter verfolgt werden soll. Ein Beispiel für einen Verzweigungspunkt ist eine Straßenkreuzung. Verzweigungspunkte sind Gebiete, in denen eine Person eine Aktion auswählt, durch die er einen Pfad fortsetzen möchte. Ein Verzweigungspunkt ist kein Punkt im Sinne einer euklidischen Definition. Vielmehr repräsentiert ein Verzweigungspunkt ein Gebiet, in dem die Bewegungsrichtung verändert werden kann.

Routenwissen wird häufig als linearer Modus räumlichen Wissens bezeichnet (vgl. [All81b], [HLP78], [PIS60]), da es eine Sequenz von geordneten, zeitlich-räumlichen Beziehungen zwischen einem Akteur und seiner Umgebung expliziert (vgl. [ASR78], [BW76], [MC82], [AK85]). Aus Gründen der kognitiven Ökonomie wird postuliert, den kontinuierlichen Fluß zeitlich-räumlicher Erfahrung während des Erlernens einer Route in Segmente (chunks) aufzuteilen (vgl. [AK85], [All81a], [All87]). Das Segmentieren von Routen ist konzeptuell eine Kategorisierung von symbolischen räumlichen Ordnungen. Die Einteilung in Segmente hat Auswirkungen auf das Schätzen von Entfernungen. Empirische Studien zeigen einen positiven Zusammenhang zwischen

geschätzten und tatsächlichen Entfernungen innerhalb von Segmenten, sowie schwach positive bzw. keine Zusammenhänge zwischen Segmenten (vgl. [AK85]). Mit Routenwissen sind weiterhin Fähigkeiten verbunden, einen Reiseplan vorab auszuwählen, in eine propositionale Struktur zu bringen und Heuristiken zu bestimmen, mit denen dieser Plan in Aktionen umgesetzt werden kann (vgl. [GBL85], [MC82]).

Der Aufbau von Routenwissen setzt voraus, Information über Lokationen zu ordnen, zu linearisieren, zu Segmenten zusammenzufassen und mit angrenzenden Segmenten verbinden zu können. Dazu werden Fähigkeiten vorausgesetzt, mit denen Richtungen und Orientierungen relativ zu vorhergehenden und nachfolgenden Segmenten bestimmt werden (vgl. [Gol91]). Ebenso wie bei Landmarkenwissen gibt es Evidenzen, daß auch Routenwissen hierarchisiert ist (vgl. [Pai70]).

Ein Minimum an Routenwissen repräsentiert eine sequentielle Abfolge von Lokationen und Distanzen zwischen Lokationen, die ein Akteur von der Startlokation bis zur Ziellokation perzipiert, sowie einer Repräsentation der Aktionen an Verzweigungspunkten (vgl. [Kui78], [Lei87], [SL88]). *Erweitertes Routenwissen* enthält Information über sekundäre und tertiäre Landmarken entlang des Pfades, Distanzen zwischen Landmarken, Relationen zwischen Landmarken und Pfadabschnitten zu einem globalen Referenzsystem (vgl. [Gol91]).

Bei der Transformation von Pfadabschnittsrepräsentationen in räumlichen Langzeitrepräsentationen kommt es zu charakteristischen Verzerrungen und Fehlern (z.B. [Tve81], [Tve92]). Die Segmentierung eines Weges wird um so effizienter durchgeführt, je differenzierter die betreffende Umgebung ist (vgl. [All81b]), wobei die Eigenbewegung einen starken Einfluß auf das Erlernen räumlicher Eigenschaften besitzt (vgl. [Acr88], [LPN81]).

### 2.3.6 Konfigurationswissen

Die dritte Art räumlichen Wissens ist *Konfigurationswissen* (vgl. [SW75]). Dieses entsteht, wenn Distanzinformation präzisiert wird, sowie Vorstellungen von Winkelzusammenhängen, Richtungen, Kontinuitäten und übergreifende Relationen abgeleitet werden. Dies erfolgt durch Integration von Routensegmenten zu höheren Repräsentationen räumlichen Wissens (vgl. [HMP76], [LSG84], [SW75]).

Konfigurationswissen kann flexibler als Landmarken- und Routenwissen verwendet werden, da Relationen zwischen Landmarken direkt zugreifbar sind. Abkürzungen und Alternativpfade können dadurch einfach bestimmt werden (vgl. [THR82], [LG83]). Insbesondere für Konfigurationswissen gibt es Evidenzen, daß der Zugriff bzw. die Repräsentation hierarchisiert erfolgt (vgl. [HJ85], [McN86], [Sho87]). Konfigurationswissen soll hier nicht eingehender diskutiert werden, da es Bestandteil des Langzeitgedächtnisses ist, welches in dieser Arbeit nicht berücksichtigt wird.

Landmarken-, Routen- und Konfigurationswissen sind die Konzepte, mit denen räumliches Wissen konzeptualisiert wird. Diese verschiedenen Wissensarten werden im Gedächtnis in Kurz- und Langzeitrepräsentationen integriert gespeichert.

### 2.3.7 Theorien zu Kognitiven Karten

Für die Gesamtheit individuellen, räumlichen Wissens im Langzeitgedächtnis werden verschiedene Begriffe verwendet: „imagery map“ (vgl. [Tro13]), „Feldkarte“ oder „Kognitive Karte“ (vgl. [Tol48], [BMB70], [DS73]), „mentale Karte“ (vgl. [Hal55]), „Schema“ und „topographische Repräsentation“ (vgl. [She62]). Der Begriff der Kognitiven Karte hat sich weithin durchgesetzt. Das Konzept der Kognitiven Karte expliziert raumbezogene Langzeitrepräsentationen, die als Bindeglied sowohl zwischen Perzeptionen, Umgebungen, als auch raumbezogener sprachlicher Beschreibungen fungieren (vgl. [Tve93], [MHH92b]).

Eine Kognitive Karte integriert Landmarken-, Routen- und Konfigurationswissen, die es einem Akteur erlauben, mentale Vorstellungen und Inspektionen über räumliche Zusammenhänge durchzuführen. Das strukturierende Prinzip der Hierarchisierung in den einzelnen Wissensformaten findet sich auch auf der Ebene der Kognitiven Karten wieder (vgl. [Tve81], [LH87], [Tve92], [MHH92a]), formale Modellierungen finden sich z.B. in [Kos87], [HHD91]). Bei der Verwendung von räumlichem Wissen kommt es zu systematischen Verzerrungen und zu Fragmentierungen der gespeicherten Wissensstrukturen (vgl. [Tve93]). Typische Fehler sind Distanzverzerrungen (z.B. [SBS80], [HM82], [KPF74], [NL82], [SS80], [Tho81]), Gruppierungsfehler (vgl. [Tve81]), Rotationsfehler (vgl. [Tve81], [Cha82], [Llo89], [LH87]) und Linearisierungsfehler (vgl. [MJ76], [SC78], [HM92], [MB83], [SM89]).

Einer der Effekte, die bei der Verwendung räumlichen Langzeitwissens auftreten, sind systematische Fehler bei der Angabe von Richtungen bezüglich Umgebungen (vgl. [SC78]). Landmarken werden räumlichen Gebiete zugeordnet, wodurch Konglomerate und Hierarchien entstehen. Reaktionszeiten bezüglich Richtungsentscheidungs- und anderen räumlichen Aufgaben über Konglomerate hinweg sind kürzer als innerhalb eines Konglomerats (vgl. [Wil75], [Mak81]). Beispielsweise lassen sich bezüglich städtischen und staatenbezogenen Umgebungen Priming-Effekte nachweisen (z.B. [Ste78], [Shu84]). Wird die Schätzung der Distanzen zwischen zwei Lokationen getestet, so zeigen sich systematische Asymmetrien (vgl. [SBS80], [Cad79]). Gleiche Distanzen werden zwischen Konglomeraten länger geschätzt als innerhalb eines Konglomerats (vgl. [All81a]). Die hierarchische Struktur Kognitiver Karten repräsentiert Referenzpunkte, Wissen über Distanzen und semantische Information über Objekte von Umgebungen (vgl. [HJ85], [Tve93]). Eine Hierarchisierung räumlichen Wissens birgt jedoch die Frage in sich, ob eine Repräsentation auf einem zwei- oder dreidimensionalen euklidischen Raum basiert (vgl. [GR73], [GS78], [GRR82], [GZ73], [Ric82]). Empirische Evidenz deutet darauf hin, daß die räumliche Information über wohlbekannte Plätze durch eine zweidimensionale Struktur repräsentiert wird (vgl. [Gol87]).

Nachdem die empirischen Befunde im Kontext des räumlichen Wissens skizziert worden sind, werden im weiteren formale Aspekte des motorischen Verhaltens vorgestellt.

## 2.4 Aspekte des motorischen Verhaltens

Ein weiterer wichtiger Interaktionsmodus mit einer Umgebung eröffnet sich durch die *Motorik*. Aus formaler Sicht kann ein *physikalischer Agent*<sup>34</sup> über *Effektoren* seine Umgebung manipulieren oder sich darin fortbewegen. Im Hinblick auf sprachliche Beschreibungen einer Umgebung soll hier nur der letztere Fall diskutiert werden.

Neben anderen Interaktionsmodi bietet die Motorik eine weitere Möglichkeit, räumliches Wissen in einer Situation zu verankern. In einem der detailliertesten Modelle zur Verankerung räumlichen Wissens in motorischen Daten, unterscheidet Kuipers in seiner *räumlichen Semantik-Hierarchie* u.a. zwischen einer *Steuerungsebene* und einer *Kausalebene* (vgl. [KL88], [Kui94]). Auf der Steuerungsebene sind *Steuerungsgesetze* definiert, die es erlauben, entlang einer Trajektorie von einem Ausgangspunkt zu einem Zielpunkt zu gelangen. Dabei werden Merkmale aus der Situation derart verarbeitet, daß ein geeignetes Steuerungsgesetz für die aktuelle Struktur der Situation festgelegt werden kann bzw. verändert werden muß. Eine Abstraktion der kontinuierlichen Werte auf der Steuerungsebene wird auf der Kausalebene verwendet, um diskrete Zustände und Übergänge zwischen Zuständen zu definieren. Kuipers unterscheidet zwischen Sichten (views) und Aktionen (actions). Sichten beschreiben den perzeptuellen Eindruck des

---

<sup>34</sup>Ein physikalischer Agent zeichnet sich dadurch aus, daß er in einer Realumgebung agiert, wohingegen ein Software-Agent (softbots) in einer synthetischen Umgebung, wie z.B. in cyberspace verankert ist.

Agenten zu einem gegebenen Zeitpunkt an einer gegebenen Lokation. Aktionen repräsentieren Trajektorien, die Steuerungsgesetzen zugeordnet sind. Aktionen und Sichten werden während der Bewegung des Agenten wechselweise bestimmt, was zu *Sicht-Aktion-Sicht*-Assoziationen führt, die, entsprechend Kuipers Hypothese, als Basis für den Aufbau Kognitiver Karten fungieren (vgl. [Kui94]). Die Kausalebene enthält die qualitativen Aktionen *Travel*( $\delta$ ) und *Turn*( $\alpha$ ), wobei  $\delta$  die Distanz der Translation vorgibt und  $\alpha$  den Winkel der Rotation (vgl. [KL88]). Durch diese beiden Operatoren lassen sich Navigationen durch eine Umgebung qualitativ beschreiben. Wesentliche Voraussetzung hierzu ist der Übergang von *kontinuierlich-quantitativen* zu *diskret-qualitativen* Repräsentationen, welcher zwischen der Steuerungs- und der Kausalebene erfolgt.

Das Modell von Kuipers ist konsistent mit der zweiten und dritten Ebene der Brooksschen *Subsumptions-Hierarchie* (vgl. [Bro86]). Dabei kehrt dieses Modell die klassische Vorgehensweise in der Robotik um, die davon ausgeht, daß aus sensorischen Daten zuerst geometrische und nachfolgend topologische abgeleitet werden (z.B. [CL85], [ME85]).

Im folgenden soll das Hauptthema dieser Arbeit, nämlich allgemeine Raumbeschreibungen und insbesondere Wegbeschreibungen, eingehend diskutiert werden.

## 2.5 Interaktion von Sprache und räumlichem Wissens

Räumliches Wissen repräsentiert das Wissen eines Betrachters über Umgebungen und Situationen. Eine zentrale Frage in der Sprachforschung ist es, in welcher Weise Sprache einerseits dazu verwendet wird, räumliches Wissen aufzubauen (z.B. [JL83], [BTF92]) und andererseits in kommunikativen Kontexten zu verwenden (z.B. [Bie67], [Cla73], [Tal83], [Her86]). Im weiteren Verlauf wird ausschließlich Bezug auf den zweiten Aspekt genommen.

Die Verwendung räumlichen Wissens in sprachlichen Beschreibungen erfordert eine Koordinierung involvierter Prozesse und Repräsentationen. Da räumliches Wissen aus unterschiedlichen Informationsquellen abgeleitet wird, kann dieses nicht als Teil des Sprachsystems verstanden werden (vgl. [Bry92]). Vielmehr ist es notwendig, räumliches Wissen in sprachliche Repräsentationen umzuwandeln. Vor allem sprachwissenschaftlich motivierte Studien haben die Interaktion von räumlichem Wissen und sprachlichen Beschreibungen untersucht (vgl. Kapitel 2.5).<sup>35</sup>

Sprachliche Raumbeschreibungen *diskretisieren* und *linearisieren* den Raum. Diese Verdichtung des wahrgenommenen Raums erfordert robuste konzeptuelle und sprachliche Strukturen. Zwischen Repräsentationen auf der räumlichen und der sprachlichen Ebene besteht ein komplexer Zusammenhang, der i.a. nicht durch eine Eins-zu-Eins-Beziehung ausgedrückt werden kann. Vielmehr kann eine räumliche Anordnung nicht nur durch genau *eine* sprachliche Form angemessen expliziert werden. Andererseits wird eine sprachliche Form als Beschreibung einer beliebig großen Anzahl von räumlichen Konstellationen verwendet. Zwischen der sprachlichen und der räumlichen Ebene transformieren Funktionen die räumlichen Repräsentationen in neue, temporäre Repräsentationen. Aus diesen Repräsentationen wird nach funktionalen Kriterien geeignetes Wissen für Sprachprozesse ausgewählt (vgl. [vC93]).

Bei der Analyse sprachlicher Beschreibung räumlicher Sachverhalte werden drei grundlegende Konzepte unterschieden: *Entitäten*, *räumliche Relationen* und *Referenzsysteme*. Als erstes werden *Entitäten* festgelegt, über die eine sprachliche Aussage gemacht wird. Diese können Teil einer größeren Entität oder weitgehend eigenständig sein. Wie in Kapitel 2.3 eingeführt wurde, sind Entitäten Vorkommen in räumlichen Sachverhalten. Einem Vorkommen ist u.a. eine Lokation und eine Identität zugeordnet. In der raumsprachlichen Literatur gibt es Arbeiten,

---

<sup>35</sup>Dies hat z.T. die Folge, daß dabei entwickelte linguistisch gefärbte Konzepte zur Strukturierung räumlichen Wissens verwendet werden. Dadurch wird eine Nähe zwischen räumlichem Wissen und Sprache suggeriert, die nicht notwendig in einer solch eindeutigen Weise vorhanden ist.

die einerseits die Lokation einer Entität (z.B. [Tal83]) und andererseits die Identität, d.h. die physikalische Ausprägung (z.B. [OB83]), als primär ansehen. Entitäten werden in der deutschen Sprache durch Substantive beschrieben, die im kommunikativen Zusammenhang verschiedene Rollen übernehmen können.

Es ist nahezu unmöglich, Entitäten eines räumlichen Sachverhalts zu beschreiben, ohne gleichzeitig räumliche Relationen zu verwenden. Vielfach werden räumliche Relationen durch räumliche Präpositionen kommuniziert (vgl. [Tal83]). Unter der Annahme, daß räumliche Präpositionen einen direkten Zugriff auf die Struktureigenschaften von räumlichen Relationen erlauben, ist dieser Zusammenhang häufig diskutiert und untersucht worden (z.B. [MJL76], [Tal83], [Her86], [HT95]). In den meisten Arbeiten werden einfach strukturierte räumliche Präpositionen wie „links von“, „über“ und „nahe bei“ betrachtet, von denen angenommen wird, daß sie mittels einer Eins-zu-Eins-Beziehung in räumlichen Relationen verankert sind (z.B. [MJL76], [OB83], [Her86], [Gap94]). Unter Annahme einer dekompositionellen Semantik wird postuliert, daß komplexere Präpositionen, wie beispielsweise „rechts auf“, sich in die Einzelbedeutungen von „rechts“ und „auf“ zerlegen lassen.<sup>36</sup>

Es gibt zwei Gründe, auf ein Objekt bzw. seine Lokation *sprachlich* Bezug zu nehmen (vgl. [HG94]):

1. Der Sprecher intendiert, das *Objekt* hervorzuheben, um die Aufmerksamkeit des Hörers darauf zu lenken.
2. Der Sprecher intendiert, die *Lokation* hervorzuheben, an der sich das Objekt befindet.

Dies erfolgt auf der Basis von räumlichen Relationen, die durch geeignete sprachliche Ausdrücke, vor allem durch räumliche Präpositionen, kommuniziert werden (vgl. [LJ93], [Kle94], [RS88], [GW95]).

Es lassen sich drei Klassen von räumlichen Präpositionen unterscheiden (vgl. [Her86], [RS88], [Her96]):

1. Projektive Präpositionen
2. Topologische Präpositionen
3. Pfadbezogene Präpositionen

Eine der prominentesten Ansätze zur Formalisierung räumlicher Relationen ist der von Herskovits (vgl. [Her86]). Die Instanzen einer räumlichen Relation in einer Situation nennt Herskovits die Klasse der Lokalisationsausdrücke (localization expressions) (vgl. [Her86, S. 7]). Dabei werden ein Lokalisations- und ein Referenzobjekt zueinander in Beziehung gesetzt. Herskovits postuliert sogenannte *geometrische Beschreibungsfunktionen*, um die Beziehung zwischen räumlichen Entitäten und Beschreibungen auf geometrischer Ebene herzustellen. Auf der geometrischen Ebene ist die sogenannte *ideale Bedeutung* von räumlichen Relationen als Bindeglied definiert. Herskovits bindet die Evaluation einer räumlichen Relation an einen bestimmten Zeitpunkt  $t$ , wodurch prinzipiell ein Bezug zu dynamischen Situationen möglich ist. Räumliche Relationen werden an semantische und funktionale Aspekte durch sogenannte *sense shifts* angepaßt. Grundlegend ist eine logische Definition einer räumlichen Relation, nach der sie entweder wahr oder falsch ist. Dieser Umstand führt zu einer Problematik, wenn nicht genau entschieden werden kann, ob

---

<sup>36</sup>Ungeklärt ist dabei die Frage, zwischen welchen Relationenklassen Kombinationen möglich sind und zwischen welchen nicht. Im letzteren Fall ist es interessant zu wissen, warum eine bestimmte Kombination nicht möglich ist.

eine bestimmte räumliche Relation erfüllt ist oder nicht. Herskovits bringt das Beispiel, daß der Satz „Das Buch ist auf dem Tisch“ auch dann noch hinreichend gut erfüllt ist, wenn eine Tischdecke dazwischen ist. Um dieses formalisieren zu können, wird der sogenannte *tolerance shift* eingeführt, der die obige Relation als Eingabe verwendet und nur approximative Gültigkeit besitzt. Um auf einer geometrischen Ebene solche Abweichungen von der Idealkonfiguration repräsentieren zu können, wird in verschiedenen Arbeiten das gebietskonstituierende Konzept einer *Anwendbarkeitsfunktion* (AF) verwendet (z.B. [ABHR87], [Her96], [Gap94], [OMT94]). Dabei wird eine räumliche Konfiguration zweier Objekte bezüglich einer räumlichen Relation auf einen numerischen *Anwendbarkeitsgrad* (AG) abbildet. Aus der Perspektive der Sprachgenerierung bestimmt der Anwendbarkeitsgrad, wie gut eine dieser räumlichen Relation  $R$  zugeordnete sprachliche Beschreibung die Beziehung zwischen einem Lokalisationsobjekt und einem Referenzobjekt<sup>37</sup> expliziert. Da sich Lagebeziehungen in dynamischen Umgebungen ändern, ist der Anwendbarkeitsgrad an einen bestimmten Zeitpunkt  $t$  gebunden (vgl. [Her86]). Die Anwendbarkeitsfunktion  $AF_R$  ermöglicht eine lokationsabhängige Bestimmung des Anwendbarkeitswertes einer räumlichen Relation  $R$  (vgl. [ABHR87], [Gap94], [Her96]). Jede räumliche Relation besitzt eine charakteristische Anwendbarkeitsfunktion. Über diese Anwendbarkeitsfunktion wird jeder Lokation des Raumes um das Referenzobjekt ein Anwendbarkeitsgrad zugeordnet. Die Anwendbarkeitsfunktion wird nicht nur in Abhängigkeit von der Distanz zum Referenzobjekt bestimmt. Distanzen zwischen den Objekten und die Winkelabweichung des zu lokalisierenden Objektes von der typischen Lokation bezüglich einer bestimmten räumlichen Relation nehmen entscheidenden Einfluß auf die Anwendbarkeit einer räumlichen Relation auf die Lage zweier Objekte (vgl. [Her86, S. 41]).<sup>38</sup>

Eine Anwendbarkeitsfunktion  $AF_R(ego, loc(x), loc(refo), t)$  ordnet jeder Lokation  $x$  im Raum um das Referenzobjekt  $refo$  zum Zeitpunkt  $t$ <sup>39</sup> in kausaler Referenz zu der Lokation des Betrachters  $ego$  einen Anwendbarkeitsgrad  $AG_R$  zu. Die Anwendbarkeitsfunktion berechnet den Anwendbarkeitsgrad in Abhängigkeit von der Distanz, der Winkelabweichung und den Größenverhältnissen zwischen dem Lokalisations- und Referenzobjekt (z.B. [Gap94], [Her96]). In den bisher entwickelten Modellen ist eine solche Strukturierung einzig auf die geometrischen Ausprägungen der Objekte bezogen und berücksichtigt nur geringfügig funktionales oder sonstiges semantisches Wissen über Objekte.

Das dritte strukturierende Element raumsprachlicher Beschreibungen sind Referenzsysteme. In der raumsprachlichen Literatur werden Referenzsysteme grundsätzlich danach unterschieden, ob sie im Sprecher oder in einem Objekt oder einer Lokation außerhalb des Sprechers verankert sind. Bezogen auf einen kommunikativen Kontext unterscheidet Clark (vgl. [Cla73]), ob ein Sprecher eine Umgebung auf sich selbst bezieht (canonical position) oder ob der Rezipient direkt gegenübersteht (canonical encounter). Hieraus leitet sich eine alternative Unterteilung ab, in der ein Sprecher oder ein externes Objekt als Zentrum eines Referenzsystems angenommen wird (z.B. [Her86]). Sprecherbezogene Referenzsysteme werden als egozentriert und objektbezogene als allozentrische Referenzsysteme bezeichnet. Wie sprachpsychologische Untersuchungen zeigen, lassen sich allozentrische Referenzsysteme weiter in solche unterteilen, die in Gegenständen und solche, die in Rezipienten verankert sind (vgl. [Sch95b]). Sprecher neigen dazu, den kognitiven Aufwand des Rezipienten für das Verstehen und Verifizieren einer raumsprachlichen Beschrei-

---

<sup>37</sup>In manchen Arbeiten wird der sprachlichen Präposition „zwischen“ eine Sonderrolle zugeordnet, in der dem Lokalisationsobjekt zwei Referenzsysteme zugeordnet werden (vgl. [ABHR87]). Diese Präposition läßt sich aber auf raumkognitiver Ebene durch die Komposition von zwei räumlichen Relationen mit jeweils einem Referenzobjekt modellieren (vgl. [ABHR87]).

<sup>38</sup>Empirisch konnten Korrelationen zwischen den Faktoren *Distanz*, *Winkelabweichung* und *Größe* der Objekte festgestellt werden (siehe [HT95]). Für projektive Relationen gibt es Evidenzen, daß die Distanz keine oder nur eine untergeordnete Rolle spielt (vgl. [Gap95a]).

<sup>39</sup>Der Einfachheit wegen wird  $t$  im weiteren Verlauf weggelassen.

bung zu verringern (vgl. [Her89], [HGH91], [Sch95b]). Dabei versetzt sich der Sprecher meist in die Position des Rezipienten, was in der vis-à-vis Anordnung einen maximalen kognitiven Aufwand bezüglich mentalen Rotationen bei der Generierung intrinsischer Beschreibungen bedeutet (vgl. [HBN87]). Dennoch werden solche rezipientenzentrierte Beschreibungen präferiert (vgl. [Meh65], [Som69]).

Zwischen Referenzsystemen und räumlichen Relationen besteht keine Eins-zu-Eins-Beziehung. Carlson-Radvansky und Irwin konnten zeigen, daß die Verwendung der räumlichen Relation “above” einfacher ist, wenn mehrere verschiedene Referenzsysteme gleichzeitig anwendbar sind (vgl. [CRI93]). Dies zeigt einen engen Zusammenhang zwischen Referenzsystemen und räumlichen Relationen, der bisher wenig betrachtet worden ist.

Wie räumliches Wissen in sprachlichen Beschreibungen, insbesondere Wegbeschreibungen, verwendet wird, ist Gegenstand der folgenden Abschnitte.

### 2.5.1 Raumdeixis

Raumsprachliche Beschreibungen lassen sich in solche unterscheiden, die sich auf das unmittelbare Umfeld des Sprechers beziehen und solche, deren Referenzen darüber hinausgehen. Referenzen auf die unmittelbare Umgebungen konstituieren das, was als *Zeigfeld* (vgl. [B34]) oder auch *Deixis* bezeichnet wird (vgl. [Kle79]). Allgemein lassen sich die Deixis des Raums und der Zeit unterscheiden (vgl. [Lev89], [Kle79]). Raumdeiktische Ausdrücke wie „hier“ und „dort“ referenzieren auf solche Entitäten, die sowohl vom Sprecher, als auch vom Rezipienten in der aktuellen Situation identifizierbar sind. Dreierlei kann ein Sprecher durch raumdeiktische Ausdrücke<sup>40</sup> vermitteln: identifizieren, informieren und bestätigen (vgl. [Fil82]). Im Hinblick auf die Verankerung von sprachlichen Ausdrücken in visuellen Daten, sind die ersten beiden Aufgaben von deiktischen Beschreibungen zu beachten. Im Kontext der Situationssemantik (vgl. [BP83]) werden solche Sprachausdrücke mit Referenten in der Situation durch Referenzfunktionen verknüpft.<sup>41</sup> Die Aufgabe des Hörers ist es, die Instantiierung der Referenzfunktion des Sprechers zu antizipieren (vgl. [Nun79]). Ein referentieller Ausdruck kann bereits vorher eingeführt worden sein, oder es besteht für den Rezipienten die Aufgabe, ein geeignetes Objekt in der Situation bestimmen. Um die Unifikation von Referenz und potentiellen Referenten zu erleichtern, ist es die Aufgabe eines kooperativen Sprechers, die Beschreibung möglichst eindeutig zu halten. Dies kann u.a. durch Angabe von visuellen Merkmalen erfolgen, die in der Situation als *salient* erscheinen. Dabei muß zwischen solchen Merkmalen unterschieden werden, die der Sprecher selbst und solche, von denen der Sprecher annimmt, daß der Hörer sie als salient empfindet. Letztendlich ist es die Aufgabe des Sprechers, eine Lokation oder ein Objekt durch geeignete Beschreibungen derart zu referenzieren, daß der Rezipient sie eindeutig identifizieren kann (vgl. [Lev89]).

Raumdeiktische Beschreibungen sind die Grundlage von verschiedenen Diskurstypen, wie z.B. Wegbeschreibungen (vgl. [Kle82]). Durch raumdeiktische Referenzen wird primär die aktuelle Position des Sprechers und Rezipienten zur Umgebung in Beziehung gesetzt, mit dem Ziel, auszuführende Aktionen eindeutig zu indizieren. Der Diskurstyp der Wegbeschreibungen wird im weiteren genauer analysiert.

<sup>40</sup>Im weiteren Verlauf kurz als deiktische Ausdrücke bezeichnet.

<sup>41</sup>In der Situationssemantik wird die Gesamtheit aller dieser Referenzen als die *Verbindungen* des Sprechers (speaker's connections) bezeichnet (vgl. [BP83]). Eine Situation im Sinne der Situationssemantik ist gleichbedeutend mit dem deiktischen Raum, der den Sprecher und den Hörer gemeinsam umgibt.

## 2.5.2 Diskursbereich der Wegbeschreibungen

Wegbeschreibungen sind ein im Alltag häufig vorkommender Diskurstyp, der sprachliche, perzeptuelle und mentale Imaginationsfähigkeiten mit räumlichem Wissen verbindet. Wegbeschreibungen sind sowohl in ihrer kommunikativen Form, als auch inhaltlich stark schematisiert (vgl. [Kle82], [Wun78]). Dies erlaubt es dem Sprecher, Information effizient auszuwählen und zu kommunizieren, solange er annimmt, daß der Rezipient über ähnliche Schemata verfügt. Solche Kommunikationsschemata reglementieren sowohl das, *was* durch eine kommunikative Handlung an Information ausgetauscht werden kann, als auch die Art und Weise, *wie* dies prototypisch erfolgt.<sup>42</sup>

Linearisierendes Element bei Wegbeschreibungen ist die Interaktion des Sprechers durch reales oder imaginiertes Verfolgen eines Weges. Wegbeschreibungen bestehen aus Referenzen bezüglich eines imaginierten oder direkt perzipierten Raums. Aus kommunikativen und kognitiven Gründen besteht die zentrale Aufgabe bei der Generierung von Wegbeschreibungen in der Auswahl solcher Referenzen, die den Weg maximal disambiguieren und andererseits den Rezipienten nicht mit zu viel Information überlasten.

Im weiteren Verlauf wird auf den Diskurstyp Wegbeschreibung eingegangen und dabei insbesondere auf die Linearisierung. Als wesentliche Aufgabe von Wegbeschreibungen werden anschließend die referentiellen Ausdrücke genauer diskutiert. Abschließend werden relevante formale Modelle vorgestellt.

### 2.5.2.1 Wegbeschreibungen als Diskurstyp

Wegbeschreibungen sind aus linguistischer Sicht vielfach untersucht worden (vgl. [Kle79], [Wun82], [Hab87], [MMP<sup>+</sup>88], [HCR90]), was auf die eindeutige kommunikative Aufgabenstellung zurückzuführen ist.<sup>43</sup> Das Ziel einer Wegbeschreibung ist es, einem Rezipienten einen Weg von einem Startort zu einem Zielort sprachlich und durch Gesten zu beschreiben. Typischerweise hat der Auskunftgebende alleine das relevante räumliche Wissen bezüglich der Problemstellung. Von einer kommunikationsorientierten Perspektive aus gehören Wegbeschreibungen zur Klasse der *Wegauskünfte* (vgl. [Kle79]). Danach besteht eine Wegauskunft aus einer *Einleitung*, einer *Zentralsequenz* (Wegbeschreibung) und einem *Abschluß*. Wunderlich und Reinelt haben die Zentralsequenz weiter in eine Phase der *Wegbeschreibung* und einer *Absicherungsphase* unterteilt (vgl. [WR82]). Diese Phasen sind in ihrer Ausführung nicht sequentiell, sondern interagieren miteinander.

Der Diskurstyp, der durch die Phase der Wegbeschreibung festgelegt ist, läßt sich in die beiden Klassen der *vollständigen* und *inkrementellen Wegbeschreibungen* untergliedern (vgl. [Maa93]). Durch eine vollständige Wegbeschreibung beschreibt der Auskunftgeber den Weg ohne direkte Referenzierung auf die beschriebene Umgebung. Der Weg wird vom Sprecher mental verfolgt. Das Ziel einer vollständigen Wegbeschreibung ist es somit, beim Rezipienten eine solche räumlich-zeitliche Vorstellung zu erzeugen, die es ihm erlaubt, den wesentlichen Informationsgehalt der

---

<sup>42</sup>Nach Herrmann und Grabowski kann das für das Sprechen erforderliche Wissen in *Was-Wissen* und *Wie-Wissen* unterteilt werden (vgl. [HG94]). Mit Was-Wissen wird der Diskurstyp charakterisiert. Um welchen Diskurstyp es sich handelt, wird entweder explizit verhandelt, ist durch die Situation vorgegeben oder wird implizit durch die Art und Weise, wie einer der beiden Kommunikationsteilnehmer das Gespräch beginnt, vorgegeben (vgl. [Sch87]). Wie-Wissen gibt vor, wie bezüglich eines Diskurstyps eine Beschreibung strukturiert wird. Soweit das Was- oder Wie-Wissen erlernt ist, werden sie zu Routinen integriert, die Herrmann et al. *Was-* und *Wie-Schemata* bezeichnen (vgl. [HG94], für empirische Untersuchungen siehe [HKDD92]). Beispiele für die Verwendung von solch schematisiertem Sprachwissen sind Wohnungsbeschreibungen (vgl. [LL75]).

<sup>43</sup>In manchen Arbeiten werden anstelle von „Wegbeschreibung“ auch andere Begriffe wie „Routeninformation“ oder „Wegauskunft“ verwendet, die aber im wesentlichen synonym verwendet werden können.

Beschreibung zu memorieren und später eindeutig mit der Real-Umgebung abzugleichen.<sup>44</sup> Drei Elemente stellt Klein als wesentliche Bestandteile einer vollständigen Wegbeschreibung heraus. Das erste Element sind *Fixpunkte* (Landmarken oder Kreuzungen). Relativ zu Fixpunkten können *Richtungen* und *Aktionen* referenziert werden. Die Aufgabe von Wegbeschreibungen liegt darin, beim Hörer einen deiktischen Raum zu erzeugen, der mit dem des Sprechers hinreichend übereinstimmt. Da vollständige Wegbeschreibungen inhärent dynamisch sind, d.h. daß die Position des Sprechers bzw. des Rezipienten sich permanent ändert, modifiziert sich auch der deiktische Raum in dem Sinne, daß er beim Rezipient erweitert wird. Um zu erkennen, was der Sprecher als Fixpunkt bezeichnet, stützt sich der Rezipient auf vorhandenes Wissen, visuell perzipierbare Objekte sowie deiktische und relative Beschreibungen zu bereits eingeführten Fixpunkten.

Die Grundlage vollständiger Wegbeschreibungen ist das intern<sup>45</sup> oder extern<sup>46</sup> memorierte räumliche Wissen des Sprechers über eine Umgebung. Soll ein Weg beschrieben werden, so aktiviert der Sprecher einen Ausschnitt dieses Wissens aus der *Kognitiven Karte*, welcher den Start- und Zielpunkt umfaßt. Dieser Ausschnitt wird als *Primärplan* bezeichnet. Der Primärplan wird entweder vollständig im voraus oder schrittweise entwickelt (vgl. [Kle79]). Klein postuliert, daß der Primärplan zuviel Information enthält, als daß er in seiner Vollständigkeit dem Hörer beschrieben werden kann. Demzufolge wird der Primärplan überarbeitet, d.h. es wird nur *relevante* Information ausgewählt. Das Ergebnis ist ein sogenannter *Sekundärplan*, der als direkte Eingabe für die Sprachgenerierung dient (vgl. [Kle79]).

Die Unterscheidung in einen Primär- und einen Sekundärplan scheint auf den ersten Blick plausibel. Der Primärplan ist im wesentlichen ein Ausschnitt aus einer kognitiven Karte, wohingegen der Sekundärplan nur noch relevante Information enthält. Als wichtige Bestandteile bei der Bestimmung des Primärplans und des Sekundärplans sieht Klein die Festlegung von Basisreferenzpunkten und von Koordinatensystemen und die Beschränkung des deiktischen Raumes auf einen Teilbereich der Kognitiven Karte des Sprechers an. Diese stark metaphorisch geprägte Sichtweise steht vor allem in Bezug auf die Repräsentation räumlichen Wissens in Kognitiven Karten im Widerspruch. Kognitive Karten unterscheiden sich von normalen Karten erheblich, wie z.B. in ihren hierarchischen Strukturen (vgl. [HJ85]) und durch Verzerrungen räumlicher Lagebeziehungen (siehe [Tve81], [Tve92]). Vor diesem Hintergrund ist u.a. zu überprüfen, ob ein Sekundärplan vom Sprecher nicht sofort bestimmt werden kann.

Vollständige Wegbeschreibungen sind eng mit den Problematiken der Wegfindung und des komplexen Problemlösens, insbesondere mit der dynamischen Entscheidungsfindung verknüpft. Mit einer vollständigen Wegbeschreibung sind häufig Randbedingungen verbunden, die durch die verwendete Bewegungsart (z.B. laufen, fahren oder schwimmen) und den Intentionen und Präferenzen der Kommunikationsteilnehmer verbunden sind. Zwei dieser Probleme sind das *Parkplatzproblem* und das *Umsteigeproblem* (vgl. [HCR90]). Tourismus- oder Effizienz Aspekte erweitern die Liste der relevante Faktoren in nahezu beliebiger Weise. Solche Interaktionen werden hier nicht weiter untersucht.

---

<sup>44</sup>Eine solche Art der Referenz wird von Appelt als vom Sprecher und Rezipienten "nicht-geteilte Konzeptaktivierung mit Identifikationsintention" bezeichnet (vgl. [App85b]). In diesem Fall besitzt der Rezipient kein expliziertes Wissen über die referenzierten Objekte. Zwischen Sprecher und Rezipient besteht eine Wissenslücke, die der Auskunftgebende durch die Beschreibung soweit wie möglich zu schließen sucht. Ziel ist es, daß der Rezipient in der Situation die referenzierten Objekte identifizieren und adäquat agieren kann.

<sup>45</sup>Wissen, welches im Gedächtnis gespeichert wird.

<sup>46</sup>Wissen, welches durch die Verwendung externer Medien, wie beispielsweise Karten, inferiert werden kann.

### 2.5.2.2 Linearisierung durch reale und imaginierte Wanderungen

Nach Levelt gibt es zwei grundlegende Determinanten, nach denen holistisches, räumliches Wissen über eine Domäne sequenzialisiert werden kann. Zum einen legt das *Prinzip der natürlichen Ordnung* eine *inhaltsbezogene Determinante* fest, nach der zu beschreibende Information so angeordnet wird, wie es die natürliche Ordnung des Inhalts angibt (vgl. [Lev89, S. 138]). Die *prozeßbezogene Determinante* kommt dann zum Tragen, wenn keine inhaltsbezogene Determinante anwendbar ist. Sie wird durch verschiedene Prinzipien beeinflusst, wie beispielsweise das *Prinzip der Verbundenheit* (vgl. [Lev89, S. 139ff.]).<sup>47</sup>

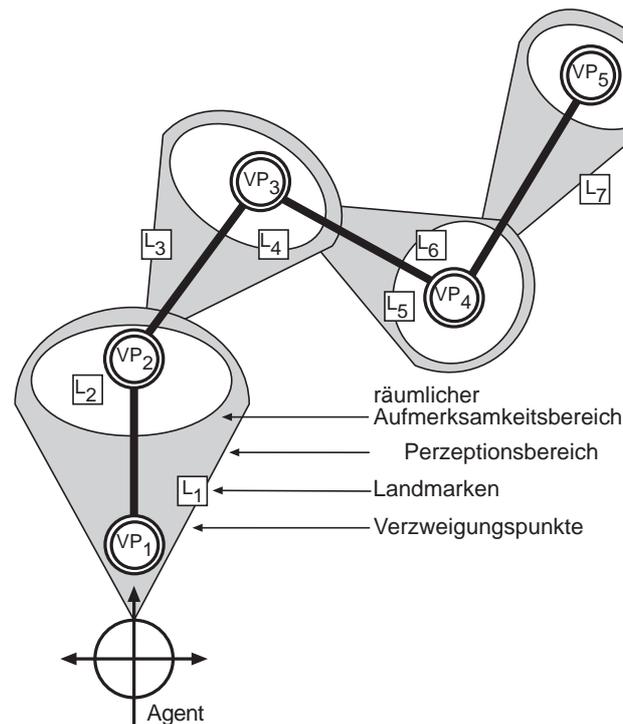


Abbildung 2.11: Linearisierung einer Wegbeschreibung durch den sich ändernden Perzeptionsbereich

Im Fall des Diskurstyps der Wegbeschreibungen ist die natürliche Ordnung durch die intendierten Trajektorien des Weges und die assoziierte Struktur der imaginierten oder direkt perzipierten Umgebung determiniert. Bei inkrementellen Wegbeschreibungen perzipiert der Auskunftgeber zu einem Zeitpunkt jeweils nur eine Situation. In dieser Situation wird eine begrenzte Menge an Landmarken ausgewählt, die entsprechend ihrer Bedeutung für die Beschreibung in die sprachliche Inhaltsstruktur integriert werden. Dazu werden solche Landmarken ausgewählt, die, ähnlich dem Ankerpunktprinzip (vgl. [CGGT87]), zu dem aktuellen Verzweigungspunkt in Beziehung stehen (siehe Abbildung 2.11). Die Linearisierung erfolgt somit entsprechend der sequentiellen Abfolge von Verzweigungspunkten entlang eines Pfades. Bezüglich Verzweigungspunkt VP<sub>2</sub> perzipiert der Agent die Landmarken L<sub>1</sub> und L<sub>2</sub>, d.h., daß beide im Perzeptionsbereich liegen. Im

<sup>47</sup>Nach Levelt werden Wohnungsbeschreibungen durch prozeßbezogene Determinanten und Wegbeschreibungen durch inhaltsbezogene Determinanten sequenzialisiert. Er führt dies darauf zurück, daß bei Wohnungsbeschreibungen keine natürliche Ordnung vorzufinden ist. Dies kontrastiert mit dem Umstand, daß Wohnungsbeschreibungen typischerweise an der Eingangstür starten und dann kreisförmig die Räume durchlaufen. Das Sequenzialisierungsprinzip ist einzig davon abhängig, ob der Durchlauf links- oder rechtsorientiert ist. Was Wohnungs- und Wegbeschreibungen gemein haben, ist die Verbundenheit der Wegabschnitte. Somit ist zu vermuten, daß Wohnungsbeschreibungen in der gleichen Weise wie Wegbeschreibungen inhaltsstrukturiert sind.

Perzeptionsbereich fokussiert der Sprecher auf einen Bereich, der um den nächsten Verzweigungspunkt liegt (räumlicher Aufmerksamkeitsbereich). Der Sprecher präferiert solche Landmarken, die im räumlichen Aufmerksamkeitsbereich liegen, da sie eine räumliche Nähe Verzweigungspunkt besitzen. Demzufolge wird nur  $L_2$  ausgewählt, da  $L_1$  von  $VP_2$  zu weit entfernt ist. Dies illustriert, daß die Beschreibung des Wegabschnitts von  $VP_1$  nach  $VP_2$  alleine durch eine lokale Beschreibung einer Aktion am Verzweigungspunkt angezeigt wird. Somit reduziert sich das Linearisierungsproblem im Kontext von Wegbeschreibungen auf ein solches für die Beschreibung einer Aktion an einem Verzweigungspunkt. Wie die empirischen Daten zeigen (siehe Kapitel 3), werden bei inkrementellen Wegbeschreibungen Aktionen in vielen Fällen durch Referenzen auf Landmarken und Zeitmarker disambiguiert. Beschreibungen von Aktionen, Landmarken und Zeitmarkern können im Deutschen in flexibler Weise permutiert werden, wodurch die Sequentialisierung der Beschreibung auf grammatikalische Regeln reduziert ist. Die Hierarchisierung des perzipierten Raums determiniert über das Prinzip der Verbundenheit die globale Sequentialisierung der Beschreibung (siehe Abbildung 2.11; Aktionen an  $VP_i$  werden vor solchen an  $VP_{i+1}$  beschrieben). Auf der lokalen Ebene eines Verzweigungspunktes werden sprachliche Ausdrücke entsprechend grammatikalischer Abhängigkeiten und persönlichen Präferenzen sequentialisiert.

In ähnlicher Weise wie bei inkrementellen Wegbeschreibungen erfolgt die Sequentialisierung bei vollständigen Wegbeschreibungen (vgl. [Kle79], [WR82]). Anstelle einer realen Bewegung durch eine Umgebung wird das Konzept einer *imaginären Wanderung* (vgl. [LL75], [Kle79]) oder eines *generischen Wanderers*<sup>48</sup> (vgl. [HG94]) verwendet. Entsprechend der Metapher der *imaginären Wanderung* kann sich der Auskunftgebende in seiner Kognitiven Karte in ähnlicher Weise wie in einer Realumgebung „umschauen“. Bisher sind aber erst Ansätze einer Theorie erkennbar, die beschreiben, welche Funktionen und Repräsentationen es erlauben, räumliche Information aus einer Kognitiven Karte zu extrahieren (siehe Kapitel 2.3.7). Sprachorientierte Arbeiten zu vollständigen Wegbeschreibungen basieren auf der Hypothese, daß die Strukturierung und Sequentialisierung des memorierten räumlichen Wissens auf ähnliche Weise wie beim Aufbau räumlichen Wissens via direkter Perzeption erfolgt. Aus diesem Grunde verwenden solche Ansätze auch eine Strukturierung entsprechend der sequentiellen Abfolge der Verzweigungspunkte. Wie Verzweigungspunkte und Landmarken referenziert werden, wird im weiteren diskutiert.

### 2.5.2.3 Referentielle Ausdrücke in Wegbeschreibungen

Die Qualität einer Wegbeschreibung ist um so höher, je geringer die Ambiguitäten bei gleichzeitiger Minimierung des kognitiven Verstehensaufwandes beim Hörer sind. Da der Sprecher keinen direkten Zugriff auf den kognitiven Aufwand beim Rezipienten besitzt, ist ein heuristisches Maß durch die Länge der Wegbeschreibung gegeben. Dadurch entsteht ein „trade-off“ zwischen der Disambiguierungsfähigkeit und der Länge einer Wegbeschreibung (vgl. [Gri75]).

Eine sprachliche Referenzbildung auf ein Objekt erfolgt selten durch Aufzählung aller Merkmale (vgl. [Lev89, S. 129]). Vielmehr wählt der Sprecher aus den memorierten oder perzipierten Merkmalen solche aus, die es dem Rezipienten erlauben, das Objekt eindeutig zu identifizieren. Wie psycholinguistische Studien zeigen, sind referentielle Ausdrücke nahezu immer ausreichend oder nicht-ambivalent. Jedoch enthalten sie meist redundante Elemente (für einen Überblick siehe [DP82]). Weiterhin werden referentielle Ausdrücke bevorzugt, die leichter zu verstehen sind als solche, die einfacher zu beschreiben sind (vgl. [RD92]). Reiter und Dale beschreiben ein Modell, in dem das präferierte Merkmal eines Objektes (z.B. Größe, Form und Farbe) bevorzugt wird, falls es die Zahl der referenzierbaren Objekte reduziert.

<sup>48</sup>Die Verwendung des Begriffs „Wanderung“ rührt daher, daß diese Untersuchungen meist mit Fußgängern durchgeführt worden sind. Jedoch lassen sich entsprechende Untersuchungen auch unter Verwendung anderer Bewegungsarten durchführen.

In einer idealen Situation kann der Sprecher eine Situation vollständig analysieren. In ressourcenbeschränkten Situationen ist dies i.a. nicht möglich. Clark und Wilkes-Gibbs haben hierzu drei Hauptfaktoren identifiziert, die die Auswahl eines referentiellen Ausdrucks in einer ressourcenbeschränkten Situation bestimmen (vgl. [CWG86]). Erstens können *temporale Beschränkungen* eine ideale Referenzierung unterbinden. Zweitens kann eine vollständige Referenzierung zu *komplex* sein, und drittens kann die *Ignoranz des Sprechers* bezüglich des Wissens des Rezipienten zu einer inadäquaten Beschreibung führen. In inkrementellen Wegbeschreibungen ist vor allem der erste Faktor von besonderer Bedeutung. Die beiden anderen Faktoren konnten im Korpus nicht identifiziert werden (siehe Kapitel 3).

Generell setzt die Verbalisierung eines referentiellen Ausdrucks voraus, daß der Sprecher und Rezipient hinreichend ähnliche Grundannahmen haben. Nicht immer ist dies gegeben, so daß es bei nicht erfolgreicher Interpretation des referentiellen Ausdrucks auf Seiten des Sprechers zur Aktivierung von Reparaturmechanismen kommt (vgl. [Edm93, S. 17]). In Bezug auf referentielle Ausdrücke werden Kommunikationssituationen unterschieden, in denen der Sprecher und der Rezipient gemeinsames Wissen über die referenzierten Objekte besitzen (vgl. [App85a], [AK87], [Hee91], [Sea69]) oder in denen die referenzierten Objekte im Aufmerksamkeitsbereich beider liegen (vgl. [RD92], [GS86]).<sup>49</sup>

Bei der Auswahl eines geeigneten Referenzobjektes zur Beschreibung einer Lokation spielen vor allem Salienzeffekte eine Rolle (vgl. [PC90]). Die Salienz eines Objektes ist von der Betrachtungsperspektive, der Bekanntheit und von den Zielen des Sprechers bzw. Rezipienten abhängig (vgl. [Lyn60]). Über die Perspektive sind perzeptuelle Eigenschaften, wie Identifizierbarkeit, Sichtbarkeit und Auffälligkeit von Einfluß (vgl. [Dev76]). Das Konzept der Salienz ist aber besonders durch die perzeptuelle Auffälligkeit beeinflusst, welche wiederum nur relativ zu einer Situation bestimmt werden kann.<sup>50</sup>

## 2.6 Modelle allgemeiner raumkognitiver Leistungen

Eine Vielzahl formaler Modelle sind zur Verarbeitung räumlichen Wissens entwickelt worden. Dabei werden zumeist entweder einzelne raumkognitive Leistungen isoliert oder mehrere integriert betrachtet.

### 1. Isolierte raumbezogene Leistungen

#### (a) Objektrepräsentationen

(z.B. [Bin71], [Neu89], [Koe90], bei der Verarbeitung visueller Daten [Bro81, Mar82, Rob63, Low87, HU90])

#### (b) Räumliche Relationen

(z.B. [RS88], [Gap94], [Han80], [CJ85], [ABHR87] [AK93], [Ege91], [RCC92], [Sch95a, OMT94], [Her91], [Zim93], [R94])

#### (c) Erkennen und Verarbeiten von Ereignissen

(z.B. [NN83a], [Her92], [Nag88], [PB95])

#### (d) Sprachliche Raumausdrücke

(z.B. [ABHR87], [RS92], [SBSZ87], [Pri93])

---

<sup>49</sup>Appelt nennt den ersten Fall "non-shared concept activation with identification intention" (vgl. [App85c]) und den zweiten "shared concept activation with identification intention" (vgl. [App85c]). Im Zusammenhang mit inkrementellen Wegbeschreibungen werden referentielle Ausdrücke des zweiten Typs verwendet. Vollständige Wegbeschreibungen hingegen verwenden nahezu ausschließlich solche des ersten Typs.

<sup>50</sup>Dieser Punkt wird häufig hervorgehoben, jedoch nahezu immer aufgrund seiner vermuteten Komplexität nur unzureichend, d.h. statisch behandelt (siehe z.B. [Edm93]).

## 2. Integrierte raumbezogene Leistungen

(vgl. [Kui77], [MS79], [Fur81], [EL82], [SVW85], [KL88], [GKS89], [LZ89], [Dav86], [MD84], [Yea88], [HW94])

Die erste Sichtweise betrachtet einzelne *Bestandteile* eines raumkognitiven Systems und entwickelt hierzu formale Modelle. Die zweite Sichtweise stellt die Frage, wie verschiedene raumkognitive Leistungen *integriert* werden können, um ein bestimmtes Problem in einer komplexen Umgebung zu lösen. In diesem Sinne sind beide Richtungen komplementär. Im weiteren werden Ansätze zur Lösung isolierter Leistungen nicht weiter diskutiert (hierzu vgl. [Her96]). In Bezug auf das hier vorgeschlagene Modell eines Agenten ist die zweite Kategorie von primärem Interesse.

### 2.6.1 Das TOUR Modell

In den ersten Arbeiten zur Modellierung raumkognitiver Leistungen durch Computermodelle ist meist eine Zwischenstellung zwischen isolierten und integrierten Leistungen eingenommen worden (vgl. [Kui77], [MS79], [EL82]). Kuipers Tour-Modell ist in der Lage, räumliches Wissen durch einen *Assimilationsprozeß* zu lernen und für Wegfindungszwecke zu verwenden (vgl. [Kui77], [Kui78]). Dieses Modell basiert auf entwicklungspsychologischen Theorien (vgl. [PIS60], [SW75]) und den räumlichen Konzepten von Lynch (vgl. [Lyn60]). In den ersten Arbeiten unterschied Kuipers nur zwischen *prozeduralem* und *topologischem* Wissen (vgl. [Kui78]). Im Laufe der Zeit erweiterte er dies zur "spatial semantic hierarchy", die fünf<sup>51</sup> verschiedene Wissens Ebenen unterscheidet (vgl. [Kui94]). In der sensomotorischen Ebene ist das gesamte räumliche Wissen des Agenten verankert. Die Information dieser Ebene steht in enger Beziehung zur Steuerungsebene, auf der Steuerungsgesetze zur Trajektorienverfolgung auf lokaler Merkmalsebene definiert sind. Die kontinuierliche Information der sensomotorischen und der Steuerungsebene wird auf diskreten Strukturen abgebildet, die zum einen *Sichten* und zum anderen *Aktionen* enthalten. Sichten repräsentieren Integrationen von sensorischer und aktionsorientierter Information zur Trajektorienverfolgung. Auf der Kausalebene werden zwei Sichten (z.B. V und V') und eine Aktion (A) zu Assoziationen integriert ( $\langle V, A, V' \rangle$ ). Sichten werden meist an Verzweigungspunkten generiert, wohingegen Aktionen solche sind, die ein Roboter benötigt, um von einer Lokation, die einer Sicht zugeordnet ist, zur nächsten zu gelangen. Dementsprechend korreliert dies mit dem Übergang von Landmarken- zu Routenwissen (siehe Kapitel 2.3.5). Auf der topologischen Ebene werden *Sicht-Aktion*-Assoziationen zu *Plätzen*, *Pfaden* und *Regionen* integriert, zwischen denen topologische Relationen, wie *Verbindung* und *Inklusion* etabliert werden. Dies korrespondiert zum Übergang von Routen- und Landmarkenwissen zu Konfigurationswissen (siehe 2.3.6). Erst nach dem Aufbau topologischen Wissens kann metrisches Wissen, wie Distanzen, Richtungen, Formen u.ä. abgeleitet werden.

In dem Modell, welches durch die *räumlich-semantische Hierarchie* gegeben ist, werden im Vergleich zu anderen Modellen die meisten Arten räumlichen Wissens integriert. Dennoch bleiben einige Fragen offen. Beispielsweise ist nicht geklärt, ob die Unterteilung in schnappschußartige Sichten und verbindende Aktionen auch in komplexen Situationen eine geeignete Konzeptualisierung darstellt.<sup>52</sup> Ebenso ist es möglich, metrische Information direkt von sensorischer Information abzuleiten, ohne eine topologische Repräsentation aufzubauen. Die Sequentialisierung bei der Entwicklung von Landmarken- und Routenwissen, so wie sie von Piaget vorgeschlagen

<sup>51</sup>Das in [KL88] beschriebene Modell wird um die *Steuerungsebene* erweitert.

<sup>52</sup>Obwohl dies von Kuipers nicht explizit beschrieben wird, ist anzunehmen, daß Sichten nur beim Stillstand des Roboters etabliert werden können. Dies führt in dynamischen Situationen, in denen der Roboter permanent in Bewegung bleiben muß, zu unakzeptablen Leistungen.

worden ist, findet in der psychologischen Literatur Gegenbeispiele (vgl. [Bla91]).<sup>53</sup> Eine mögliche Erweiterung besteht somit in der Auflösung dieser sequentiellen Architektur.

Das Tour-Modell ist die Grundlage für das sogenannte NX-Roboter-Modell (vgl. [KL88]). Dieses Modell erweitert und generalisiert das Tour-Modell durch Mechanismen zur Identifikation von Pfaden und Plätzen während der Navigation. Der NX Roboter kann eigenständig Pfade verfolgen.

### 2.6.2 Das Modell ELMER

In einem anderen Modell habe McCalla und Schneider einen Agenten, *ELMER*, entwickelt, der auf Pfaden durch eine synthetische, zweidimensionale Umgebung navigiert (vgl. [MS79]). Der Kern von ELMER ist ein hierarchisierter Planer. Planfragmente können von ELMER parallel ausgeführt werden, wobei durch Prioritäten Konflikte gelöst werden. Die Repräsentation der Umgebung enthält Pfade und Regionen. Besteht die Aufgabe darin, von einer Region  $b$  zu einer Region  $c$  zu gelangen, und gibt es einen Pfad, der zwei Überregionen  $B$  und  $C$ , die  $b$  bzw.  $c$  enthalten, verbindet, so verwendet der Agent diesen Weg. Ansonsten werden über Transitivitätseigenschaften Regionen rekursiv gesucht, mittels denen zwei Gebiete verbunden werden können.

### 2.6.3 Das Modell von Elliott und Lesk

Einen ähnlichen Suchalgorithmus wie in ELMER verwendet das Modell von Elliott und Lesk (vgl. [EL82]). Ausgehend von der Fragestellung, welcher Pfad für einen Menschen der "beste" ist und welcher Algorithmus den „kürzesten“ Pfad liefert, haben sie verschiedene Varianten getestet. Aus empirischen Untersuchungen ist bekannt, daß die Anzahl von Abbiegevorgängen minimiert wird, weswegen sie in diesem Modell mit zusätzlichen Kosten verbunden sind.<sup>54</sup> Basierend auf einer Straßenkarte haben Elliott und Lesk festgestellt, daß Suchalgorithmen mit Tiefensuche ca. 45% schneller sind als mit Breitensuche. Ein Test mit Probanden ergab, daß eine Kombination aus Tiefensuche mit einer hierarchisierten Suche, die zwischen Haupt- und Nebenstraßen unterscheidet, „adäquate“ Lösungen liefert.

### 2.6.4 Das Modell TRAVELLER

Das von Leiser und Zilbershatz entwickelte TRAVELLER Modell gibt eine Formalisierung für die Organisation und die Entwicklung von Routenwissen an (vgl. [LZ89]). Entsprechend empirischer Arbeiten (insbesondere [Pai70]) wird Routenwissen in einer Basisnetz- und einer Sekundärnetzebene hierarchisiert. Im TRAVELLER Modell repräsentieren sogenannte *Zentroiden* bekannte Lokationen, die der Agent früh erlernt hat und häufig erreicht (ähnlich dem allgemeinen Begriff der Landmarke (vgl. Kapitel 2.3.4) und der Ankerpunkttheorie (vgl. [CGGT87])). Die Suche von einem Knoten  $A$  zu einem Knoten  $B$  erfolgt in drei Schritten. Ist  $A$  kein Zentroid, so sucht er einen Pfad zum nächsten Zentroid. Von dort aus sucht er einen Pfad zu einem Zentroiden, der einen Pfad zu  $B$  besitzt, um im letzten Schritt von diesem Zentroiden zu  $B$  zu gelangen. Die Erkundung der 2D-Umgebung erfolgt zufällig. Pfade zwischen Knoten sind gerichtet. Da der Agent die Strategie verwendet, immer möglichst schnell zu einem Zentroiden zu gelangen, werden wenig Pfade zwischen Knoten verwendet, die nicht Zentroiden sind.

<sup>53</sup>Aus einer ingenieursorientierten Perspektive ist die räumlich-semantische Hierarchie nicht zu beanstanden, solange sie ihren Zweck erfüllt. Kuipers postuliert jedoch auch eine kognitionswissenschaftliche Relevanz seiner Arbeit, weswegen sie sich auch an dieser zu messen hat.

<sup>54</sup>Ein Rechtsabbiegevorgang „kostet“ eine 1/8 Meile und ein Linksabbiegevorgang eine 1/4 Meile. Für eine Suche in einem Gebiet mit Linksverkehr müssen diese Gewichte vertauscht werden.

Die Behauptung von Leiser und Zilbershatz, daß der TRAVELLER im Hinblick auf eine kognitive adäquate Repräsentation des Wissens einer kognitiven Karte bessere Eigenschaften als das TOUR-Modell besitzt, wird nicht erfüllt. Im Gegensatz zum TOUR-Modell basiert das Lernen des Routenwissens auf einer Rasterrepräsentation. Diese enthält inhärent metrische Beziehungen, von denen sich andere räumliche Relationen direkt ableiten lassen. Das topologische Konzept der räumlichen Nähe ist eine triviale Folgerung daraus. Dadurch kehrt das TRAVELLER-Modell die Entwicklung räumlichen Wissens in die Ableitung topologischem von metrischem Wissen um, was empirischen Befunden widerspricht (siehe Kapitel 2.3).

### 2.6.5 Das Modell NAVIGATOR

Das von Gopal entwickelte Modell des NAVIGATOR formalisiert Fähigkeiten, von einer Umgebung räumliches Wissen zu extrahieren und in bestehende Wissenstrukturen langfristig zu integrieren (vgl. [Gop88], [GKS89]). Bei der Konzeption des Modells sind psychologische Evidenzen bezüglich raumkognitiver Leistungen berücksichtigt worden. Entsprechend einem Agentenkonzept sind im NAVIGATOR einfache Perzeptionsfunktionen, raumkognitive und einfache motorische<sup>55</sup> Funktionen integriert werden (siehe Abbildung 2.12).

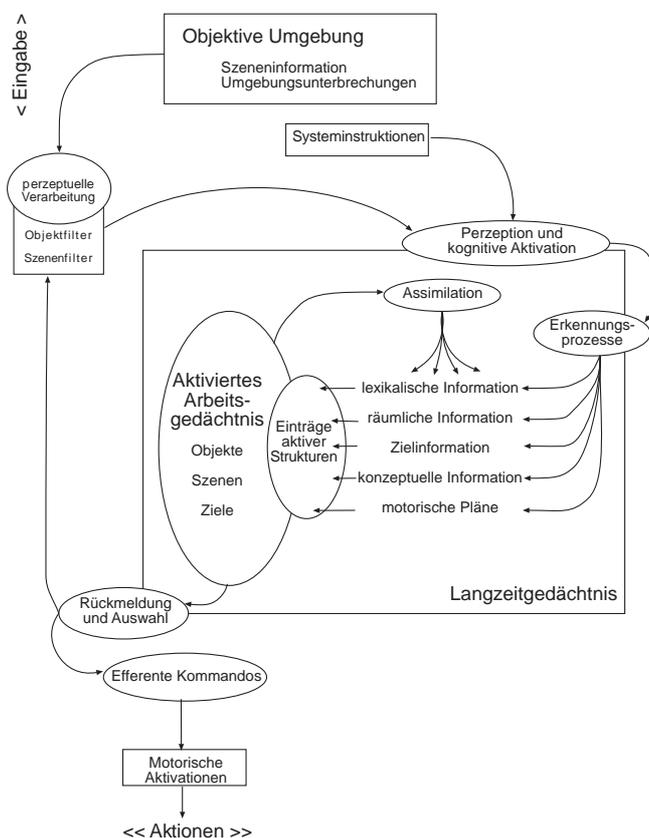


Abbildung 2.12: Schematische Darstellung der funktionalen Einheiten des NAVIGATOR-Modells nach Gopal (vgl. [GKS89])

Das NAVIGATOR-Modell umfaßt eine zweidimensionale, rasterförmige Umgebung und das eigentliche *kognitive Modul* (NV). NV bewegt sich durch von außen vorgegebene Anweisungen von

<sup>55</sup>Es sind, ähnlich wie auf der topologischen Ebene des TOUR-Modells, nur Translationen und Rotationen um 90° Winkel möglich.

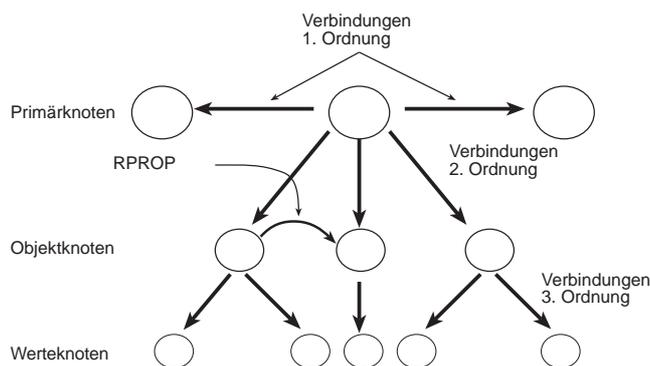


Abbildung 2.13: Räumliche Repräsentationen im NAVIGATOR-Modell

Lokation zu Lokation. In einer lokalen Umgebung bestimmt NV die Lokationen von Verzweigungspunkten (primary nodes), ausgewählten Objekten (object nodes) sowie einfachen räumlichen Relationen (RPROP link) zwischen Objekten (siehe Abbildung 2.13). Räumliche Relationen werden ausschließlich auf der Basis eines egozentrischen Referenzsystems etabliert. Ähnlich wie für das Modell von Leiser und Zilbershatz hat der Agent Zugriff auf metrisch exakte Lokationsinformation über Entitäten der lokalen Umgebung. Räumliche Relationen zwischen Objekten werden davon in stark vereinfachter Form abgeleitet. Die räumlichen Repräsentationen integrieren Landmarken- und Routenwissen. Konfigurationswissen wird nicht behandelt. Für die Auswahl von Objekten wird das Konzept der *Salienz* verwendet, welches die „perzeptuelle und subjektive Wichtigkeit“ eines Objektes bestimmt. Die „objektive“ Umgebung ist durch ein hierarchisch aufgebautes Netzwerk repräsentiert. Jedem Verzweigungspunkt sind Landmarken zugeordnet, die die Knoten auf der zweiten Ebene der Repräsentation bilden. Diese Knoten sind durch Kanten verbunden, welche räumliche Relationen zwischen Landmarken bezüglich einer bestimmten Perspektive repräsentieren. Jedem dieser Knoten ist eine Menge von Unterknoten zugeordnet, die die Werte ihrer Merkmale enthalten. NV enthält als Eingabe einen vertikalen Ausschnitt aus dieser Netzrepräsentation, der mit *Salienzbewertungen* der Landmarken angereichert ist.<sup>56</sup> Mittels Schwellwerten wird evaluiert, ob eine Landmarke eine ausreichende Salienz besitzt, um in die Repräsentation der Situation aufgenommen zu werden. Wird der typassozierte Schwellwert überschritten, so wird diese Repräsentation mit gespeichertem Wissen verglichen. Bei positivem Vergleich wird die Repräsentation in das *Arbeitsgedächtnis* geladen. In ähnlicher Weise erfolgt die Integration von Objektrepräsentationen einer *Szene* mit gespeichertem Wissen über andere Szenen.

Weiterhin verfügt der NV über Funktionen des *Vergessens*. Dazu wird der Salienzwert einer Entität im Arbeitsgedächtnis proportional zur Zeit schrittweise vermindert.<sup>57</sup> Bei wiederholter Perzeption einer Entität wird der Salienzwert erhöht.<sup>58</sup> Hat NV in einer Umgebung ausreichendes räumliches Wissen aufgebaut, ist er in der Lage, sich auf Anfrage zu einer Lokation zu bewegen, die der beschriebenen Situation am nächsten kommt.<sup>59</sup> Durch mustergesteuerten Vergleich sucht NV im Arbeits- und Langzeitgedächtnis eine geeignete Szene. Der Aufbau raumbezoge-

<sup>56</sup>Die interne Repräsentation räumlichen Wissens ist mit dem Format des externen, räumlichen Wissens identisch.

<sup>57</sup>Die entsprechende Funktion des Salienzwertes in Abhängigkeit von der Zeit lautet:  $S_t = S_0 * e^{-\beta(t-t_0)}$ , wobei  $t_0$  der Zeitpunkt der Perzeption einer Entität und  $\beta$  ein globaler Parameter ist.

<sup>58</sup>Die Erhöhung des Salienzwertes bei n-maliger Perzeption lautet:  $\delta S_n = S_0 * \alpha^{n-1}$ , mit der Konstanten  $\alpha < 1$ . Hieraus folgt, daß die Gesamtsalienz  $S_n$  eines Objektes sich wie folgt berechnet:  $S_n = \sum_{i=2}^n S_0 * \alpha^{i-1} = \frac{S_0}{1-\alpha} * (1 - \alpha^n)$ .

<sup>59</sup>Eine Beispielfrage ist: „Gehe zu einem roten Haus mit einem blauen Auto und einem hölzernen Zaun neben dem Auto.“ (vgl. [Gop88]).

ner Repräsentationen durch den NV wird dadurch stark vereinfacht, daß beide das identische Format besitzen. Woher diese Repräsentationen stammen, bleibt offen. Problematisch ist die Verwendung des Saliens-Konzepts. Einerseits leiten sich Saliens von quasi-visuellen Merkmalen ab, andererseits dienen sie als Maß der Bedeutung einer Entität. Obschon insbesondere visuelle Saliens einen Einfluß auf die Bedeutung einer Entität haben, ist eine Reduzierung des Bedeutungsbegriffs auf Saliens eine extreme Verkürzung, zumal nicht beschrieben wird, wovon Saliens abgeleitet werden. Da Salienswerte Landmarken und Verzweigungspunkten unabhängig vom Standpunkt von NV statisch zugeordnet sind, ist das perzeptuelle Modul eher eine Absichtserklärung als eine Modellierung.

Die Unterscheidung in eine *objektive Umgebung* und ein *kognitives Modul* erscheint für die Untersuchung raumbezogener kognitiver Fähigkeiten als geeigneter Ansatz, um die Dynamik der involvierten Prozesse und Repräsentationen besser untersuchen zu können. Letztendlich stellt das Modell einen interessanten Versuch der Modellierung des Prozesses des räumlichen Wissenserwerbs dar.

### 2.6.6 Das *Basic Agent*-Modell

Das von Vere und Bickmore entwickelte Modell eines *Basic Agent*, mit Namen *Homer*, integriert Funktionalitäten der NL-Generierung- und Analyse, des zeitlichen Planens, des Schlußfolgerns, der Planausführung, der simulierten symbolischen visuellen Wahrnehmung und eines episodischen Gedächtnisses (vgl. [VB90]). Den Kern bildet das zeitliche Planungssystem, welches auf dem Modell *DEVISER V* basiert (vgl. [Ver85]). Der Agent bewegt sich in einer simulierten 2D-Meeresumgebung. Er kann sprachliche Beschreibungen bezüglich dieser Umgebung analysieren und generieren. Ferner verfügt er über einen „visuellen“ Aufmerksamkeitsbereich, in dem er statische und bewegte Objekte erkennen und zur Repräsentation der Umgebung hinzufügen kann. *Homer* verfügt über ein episodisches Gedächtnis, mittels welchem er a posteriori Schlußfolgerungen über bereits ausgeführte Aktionen durchführen kann. Grundlegend ist hierfür ein lineares Zeitmodell.

Räumliche Beschreibungen werden relativ zu einem egozentrischen Referenzsystem des Agenten generiert bzw. interpretiert. Die Bedeutung von Verben ist durch einen *Frame*-Ansatz realisiert, in dem Aktionsverben durch sogenannte *state transition semantics* repräsentiert sind. Räumliche Sprachelemente, wie beispielsweise Präpositionen, werden in dem vorgeschlagenen Modell nur implizit formalisiert. Ebenso verfügt der Agent über kein explizites Konzept räumlicher Referenzsysteme, weswegen Raumbeschreibungen nur im deiktischen Gebrauch generiert und analysiert werden können. Weiterhin agiert die Implementation nicht in Realzeit, was an den aufwendigen Planungs- und Schlußfolgerungsmechanismen liegt.

Zusammenfassend läßt sich sagen, daß der Beitrag, der durch das Modell erbracht worden ist, in der Integration komplexer Leistungen besteht. Der Schwerpunkt liegt in *Homer* auf den Planungs- und Schlußfolgerungsmechanismen und der Anbindung an NL-Mechanismen.

### 2.6.7 Das Modell SPAM

Das von McDermott entwickelte SPAM<sup>60</sup>-Programm ist als autonome Wissensbasisverwaltung für räumliches Wissen konzipiert worden (vgl. [MD84]). SPAM besteht aus drei Modulen. Die *Integrationseinheit* integriert räumliches Wissen in einer Kognitiven Karte, durch welche per Anfrage über zwei Anfrageeinheiten auf einzelne oder eine Liste von Objekten zugegriffen werden kann. Zur Lokation eines Objektes ist SPAM ein Referenzsystem zugeordnet. Relationen

---

<sup>60</sup>spatial module

zwischen Objekten sind über Relationen zwischen Referenzsystemen definiert. Eine Art der Relation besteht in der Transformation der Koordinaten des einen in die des anderen. Eine andere Relationenart betrifft die Winkeldifferenzen der Orientierungen zweier Referenzsysteme und eine weitere den Unterschied der Größenverhältnisse. Bedingt durch die Verwendung von SPAM im Kontext der Verarbeitung visueller Daten sind die Grundelemente, ähnlich wie bei anderen Modellen (vgl. [Mar82], [Bin71], [Bro81]), erweiterte Zylinder mit sphärischen Enden, sogenannte "circyls" (vgl. [MD84]). Mit diesen Konzepten können Relationen zwischen Teilvolumina eines Objektes repräsentiert werden.

SPAM enthält eine Reihe von Problemen. Beispielsweise sind Objektgrenzen nicht explizit repräsentiert. Dadurch lassen sich keine qualitative räumliche Relationen definieren. Das größte Problem von SPAM ist, daß es unabhängig von einer Domäne entwickelt wurde und dadurch auf Funktionen beruht, die eine zu große Komplexität besitzen.

### 2.6.8 Das Modell MERCATOR

Das MERCATOR-Modell baut durch Erkundung einer Umgebung eine Kognitive Karte auf (vgl. [Dav86]). In diesem Modell werden drei Wissensstrukturen unterschieden:

1. Wissen über Dimensionen, Orientierungen und Winkel innerhalb eines Objektes sowie zwischen Objekten
2. Wissen über Objektformen
3. Wissen über die Größe eines Objektes

Das MERCATOR-Modell steht in direkter Nachfolge von McDermotts SPAM, weswegen der Fokus ebenso wie dort, auf visuellen Funktionen und Beschränkungen gerichtet ist. Die zweidimensionale, raumbezogene Repräsentation gliedert sich in eine Ebene, auf der Koordinaten und Kanten die Ausrichtung und Länge von Teilen eines Objektes repräsentieren. Auf der zweiten Ebene wird die Güte der durch Kanten und Knoten approximierten Form mit der der tatsächlichen Objektformen, entsprechend einer Toleranzmetrik, verglichen. Diese Strukturen werden einerseits verwendet, um visuelle Eingaben zu repräsentieren und in der Kognitiven Karte abzuspeichern und andererseits, um neue Eingaben mit den Repräsentationen der Kognitiven Karte zu vergleichen. Das Modell ist somit darauf ausgerichtet, immer eine Situation als Ganzes zu analysieren, was mit wachsender Komplexität der Situation immer schwieriger wird. Ein Objekt-konzept ist nicht vorgegeben.

### 2.6.9 Das Modell von Yeap

Eine ähnliche Fragestellung wie im MERCATOR-Modell liegt dem Modell von Yeap zugrunde (vgl. [Yea88]). Ausgehend von visuellen Daten einer Wohnungsumgebung, formalisiert er einen Prozeß zur Erstellung einer Kognitiven Karte (CMP<sup>61</sup>). Der CMP besteht aus zwei Modulen, die eine *grobe* (RCM: raw cognitive map) und eine *vollständige Kognitive Karte* (FCM: full cognitive map) erstellen. Die RCM bildet, in Anlehnung an die 2 1/2-D Repräsentation des Marrschen Modells (vgl. [Mar82]), visuelle Daten direkt auf Orientierungen und zwei-dimensionalen Kantenbeschreibungen von Objekten relativ zu einem globalen Referenzsystem ab. Eine gesamte Situation, relativ zu einer konstanten Perspektive, definiert eine sogenannte *absolute Raumrepräsentation* (ASR: absolute spatial representation). Bewegt sich der Agent durch den Raum,

---

<sup>61</sup>cognitive map process

werden mehrere ASRs zu einer *relativen Raumrepräsentation* (RSR: relative spatial representation) sequenzialisiert. Durch die Verkettung von ASRs werden Pfade und Distanzen implizit in der Ausdehnung eines einzelnen ASR repräsentiert. Yeap postuliert eine Korrespondenz einer Sequenz von ASRs mit Routenwissen und dem räumlichen Wissen innerhalb eines ASR mit Konfigurationswissen. Offen bleibt die Frage, wie aus einer RSR Konfigurationswissen abgeleitet werden kann. Ähnlich dem TOUR-Modell wird ein Übergang von einer kontinuierlichen (ASR) zu einer diskreten raumbezogenen Repräsentation (RSR) verwendet, um topologisches Wissen zu etablieren. Das Modell basiert auf der Annahme einer statischen Umgebung, in der sich einzig der Agent bewegt. Die verwendeten räumlichen Relationen repräsentieren ausschließlich Kontaktrelationen zwischen Flächen, wodurch sie als Grundlage für sprachliche Prozesse nicht geeignet sind. Da Routen- und Landmarkenwissen in diesem Modell nur implizit repräsentiert sind, erscheint es als Grundlage für sprachliche Beschreibungen als nicht ausreichend.

### 2.6.10 Zusammenfassung raumverarbeitender Modelle

Die hier vorgestellten Modelle raumkognitiver Leistungen geben einen Querschnitt, in welcher Weise räumliches Wissen allgemein formalisiert wird. Zu unterscheiden sind Modelle, die ein Agentenkonzept verwenden, von solchen, die isolierte Leistungen untersuchen (vgl. Tabelle 2.3). Das TOUR-Modell ist eines der vollständigsten Modelle, da es einerseits eine Vielzahl räumlicher Wissensarten integriert und andererseits als Grundlage für die Implementation eines physikalischen Agenten verwendet wurde. Vernachlässigt werden in diesen Modellen inwieweit Ressourcenbeschränkungen das Verhalten verändern. Es wird stets davon ausgegangen, daß es keinerlei zeitliche Beschränkungen gibt. Gleichfalls werden Referenzsysteme nur implizit behandelt. In den meisten Modellen werden allozentrische Referenzsysteme verwendet.

Im folgenden wird die Unterklasse raumverarbeitender Modelle diskutiert, die Funktionalitäten zur Generierung von Wegbeschreibungen formalisieren.

Modelle	TOUR	ELMER	SOCCER	Elliott	NX	TRAVELLER	NAVIGATOR	Basic Agent	MERCATOR	SPAM	Yeap	MOSES <sup>a</sup>
Agentenkonzept	○	○			●	○	●	●	●		○	●
2D-Umgebungen	●	●	●	●	●	●	●	●	●		●	●
3D-Umgebungen					●					●		●
Realumgebungen			●		●							
Geometrische Repräsentation	○	●	●	●		●	●	●	●	●	●	●
Visuelle Datenverarbeitung			●				○	○				●
Berechnung visueller Salienzen				●			●					●
Natürlichsprachliche Eingabe								●				●
Natürlichsprachliche Ausgabe			●					●				●
adaptiv							○	●				●
“anytime” Verhalten								○				●
Multiple Objektrepräsentationen			●						●	●		●
Topologische Relationen	●	●	●		●							○
Ordnungsrelationen			●				●	●				●
Metrische Relationen	●		●	○	●	●		●	●	●	●	●
Landmarkenwissen	●				●		●	●	●			●
Routenwissen	●	●	●		●	●	●	○			○	●
Konfigurationswissen	●				●	○		○			○	○
Egozentrische Referenzsysteme							●	●				●
Allozentrische Referenzsysteme	●	●	●		●	●	●				●	●
Globale Referenzsysteme			●						●	●		●
Kognitive Karten	●	●			●	●	●		●	●	●	

<sup>a</sup>Name des implementierten Agenten des in Kapitel 4 beschriebenen funktionalen Modells

Tabelle 2.3: Systemvergleich: ● indiziert Modellierung entsprechender Funktionalitäten; ○ indiziert partielle Modellierung

## 2.7 Modelle zur Generierung von Wegbeschreibungen

Zur sprachlichen Analyse von Wegbeschreibungen gibt es empirische Studien, theoretische (vgl. [Hab87], [HK91], [HCR90], [Cou95]) und implementierte Berechnungsmodelle (vgl. [RS87], [MMP<sup>+</sup>88], [Mül88], [Car91], [Dav89]). Diese setzen sich ausschließlich mit vollständigen Wegbeschreibungen auseinander. Theoretische Modelle sind konzeptuelle Beiträge, die hauptsächlich Funktionen identifizieren, von denen angenommen wird, daß sie bei der Generierung von Wegbeschreibungen eine Rolle spielen. Implementierte Modelle legen den Schwerpunkt auf die Operationalisierbarkeit und Implementation solcher Funktionen.<sup>62</sup>

### 2.7.1 Das formale Modell von Habel

Habel geht, wie bei einer Reihe der nachfolgenden Modelle, von einer Trennung in einen Primär- und einen Sekundärplan aus (vgl. [Hab87]), welche auf die Arbeit von Klein zurückgeht (vgl. [Kle79]). Der Gesamtprozeß untergliedert sich in Prozesse der *Wegsuche*, der *Routengenerierung* und der *Generierung natürlich-sprachlicher Routenbeschreibung* (vgl. Abbildung 2.14). Um zu unterscheiden, was wesentlich für die Wegbeschreibung ist, muß der Primärplan vollständig überarbeitet werden. Ein solcher Prozeß wandelt einen *Weg* in eine *Route* um. Hierbei werden Wegabschnitte zusammengefaßt, aber auch segmentiert und als Verzweigungspunkte „erkannt“. Aufgrund der Prämisse, ein kognitiv-adäquates Modell zu entwickeln, kann eine Richtungs- und Landmarken-Information nicht einfach an Wegabschnitte „angehängt“ werden. Um eine Trennung von geographischer Weginformation und Information über räumliche Relationen und Landmarken zu erreichen, führt Habel sogenannte *Referentielle Netze* ein (vgl. [Hab86]). Knoten und Kanten eines Referentiellen Netzes korrespondieren zu Knoten und Kanten des Wegenetzes. Die Erweiterung, die im Zusammenhang mit Wegbeschreibungen interessant ist, bezieht sich auf die Struktur der Knoten. In einem Knoten sind Zusatzinformationen abgelegt, die zum einen Attribute und zum anderen Beschreibungen, sogenannte *Designationen*, enthalten. Ein *Platz* oder ein *Knoten* zu sein, ist ein Attribut, wohingegen räumliche Lage-, Teil-von-, Bestandteilsbeziehungen und Namensbezeichner durch Designationen repräsentiert werden. Entsprechend der *dual-coding*-Theorie (vgl. [Pai71]) ist eine propositionale Repräsentationsform nicht ausreichend, um Richtungs- und Raumlagekonzepte einzuführen (vgl. [Hab87], [Pri93]). Aus diesem Grund verwendet Habel *Zellmatrizen* als „bildhaftes“ Repräsentationsformat.<sup>63</sup> Ein Ausschnitt eines Wegenetzes wird hierzu in einem matrizenartigen Netz von aneinanderliegenden Zellen unterteilt. Jeder Zelle wird entweder genau ein Straßenelement, eine Landmarke oder ein Nullwert zugeordnet. Diese Zellrepräsentation hat laut Habel einen Bezug zur depiktionalen Repräsentation im Sinne von Kosslyn (vgl. [Kos80], [Kos83]). Zellmatrizen repräsentieren laut Habel topologische Beziehungen, die zur Bestimmung von räumlichen Relationen um metrische Konzepte wie Distanz und Winkel erweitert werden. Mit Hilfe der Zellmatrizen lassen sich Gebiete zwischen Objekten repräsentieren, die es erlauben, die Bedeutung von „vor“, „links“, „rechts“ und „zwischen“ ansatzweise zu bestimmen. Um die Lokation eines Objekts mit Hilfe von referentiellen Netzen bestimmen zu können, gibt Habel folgende Schritte an. Zuerst muß das Objekt in einer Depiktion lokalisiert werden, um Nachbarschaftsbeziehungen etablieren zu können. Im nächsten Schritt wird ein Gebiet ausgewählt, in dem sich die relevanten Objekte und die aktuelle Lokation des Betrachters befindet. Zuletzt bedarf es der sprachlichen Beschreibung. Habel gibt keine

<sup>62</sup>Es wird nur eine Auswahl von Modellen diskutiert, da sich andere Modelle auf diese zurückführen lassen (wie z.B. [Mar89], [Gry92], [May92], [BB93]).

<sup>63</sup>Eine genauere Analyse der Zellmatrix-Repräsentation ergibt, daß es sich nur um graphisch dargestellte propositionale Beziehungen handelt. Eine Zelle kann durch einen Knoten repräsentiert werden, der bis zu vier Relationen zu Nachbarknoten enthält. Jeder Knoten enthält ansonsten die gleiche Information wie Zellen in der depiktionalen Darstellung. Somit ist ein Zellmatrizenformat kein wirklich depiktionales Format.

Operationalisierung dieser einzelnen Schritte an.

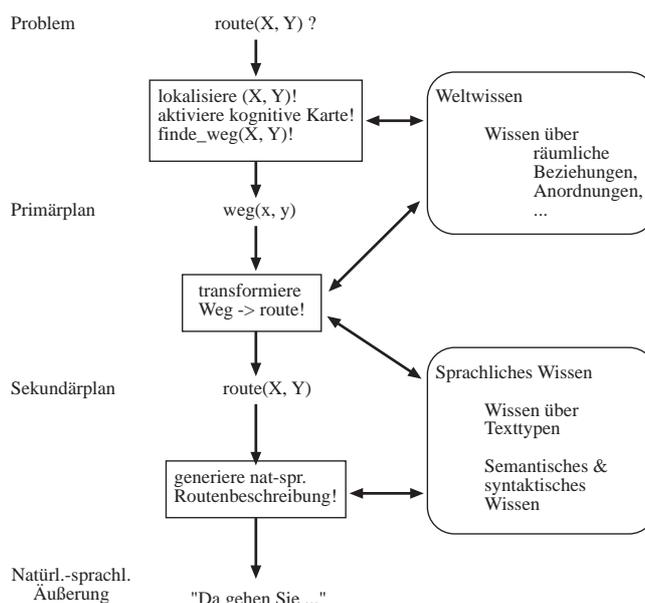


Abbildung 2.14: Interne Struktur eines Routenauskunftssystems nach Habel (vgl. [Hab87])

Ein Referentielles Netz ist eine formale Repräsentation einer Kognitiven Karte. Dieses ist Grundlage für das Generierungsverfahren für Routenbeschreibungen. Auf drei Regeln basiert die Erstellung einer *guten* Wegbeschreibung. Erstens erfolgt die Segmentierung von Wegen zu Routen an Verzweigungspunkten. Zweitens werden Verzweigungspunkte durch Landmarken beschrieben und Richtungsänderungen häufig an Landmarken orientiert. Drittens werden längere Wegabschnitte durch unterstützende Beschreibungen indiziert. Die Routenrepräsentation dient als Eingabe für den Sprachgenerierungsprozeß. Habel unterscheidet zwischen *echten* Wegbeschreibungen und Wegbeschreibungen mit *Handlungsanweisungen*. Echte Wegbeschreibungen beziehen sich „ausschließlich auf Objekte und ihre Lage bzw. Lagebeziehungen“ (vgl. [Hab87]), wohingegen im zweiten Fall Handlungsanweisungen expliziert werden. Das Ziel dieser Unterteilung ist es, sich von den rein aktionsgetriebenen Beschreibungen abzusetzen. Ein Beispiel für eine „reine“ Lagebeziehung ist die folgende:

„... da, wo das Cafe Bley ist, wo man draußen sitzen kann ...“ (vgl. [Hab87, S. 37])

Lagebeschreibungen dienen in vollständigen Wegbeschreibungen durch Einführung von Referenzobjekten dem Aufbau von Sekundärdeixen (vgl. [Kle79]). Da die rein geometrische Repräsentation von Objekten durch Koordinaten in einem Koordinatensystem nur unzureichend einer kognitiv-adäquaten Modellierung gerecht wird, gibt Habel an, daß hierfür geometrisch exakte Konzepte wie Distanz und Winkel durch vage Konzepte ersetzt werden müssen. Ebenso stellt er die Granularität von *Kognitiven Karten* zur Diskussion, die durch eine hierarchische Organisation dieser repräsentiert werden kann.

### 2.7.2 Das formale Modell von Höök und Karlgren

Vor einem mehr ingenieurwissenschaftlichen Hintergrund untersuchen Höök und Karlgren die Kommunikationsstruktur von Wegbeschreibungen (vgl. [HK91]). Dazu unterscheiden sie zwei Gruppen von Rezipienten: *Touristen* und *Stadteinwohner*. Durch empirische Untersuchungen

gestützt, in denen sie Touristen und Einwohner befragten, stießen sie auf einen *Segmentierungsprozeß* bei der Generierung von Wegbeschreibungen (siehe hierzu Kapitel 2.3.5), welcher einerseits von der räumlichen Struktur und andererseits vom *Wegtyp* abhängt. Sie klassifizieren Pfadelemente mittels einer *Hierarchie von Wegetypen* (Autobahnen, große Straßen, normale Straßen und kleine Straßen)<sup>64</sup>. Wegbeschreibungen, wie sie von Touristen und Einwohnern gegeben werden, unterscheiden sich vor allem in ihrer *Auflösung* bezüglich der beschriebenen räumlichen Strukturen. Wird eine Wegbeschreibung für einen Touristen gegeben, so werden alle Wege, auffällige Verzweigungspunkte und auszuführende Aktionen detailliert beschrieben. Für Einwohner ist die Beschreibung i.a. weniger detailliert. Diese beiden Typen bilden die Extrempole eines Spektrums. Wegbeschreibungen für Touristen bezeichnen sie als „prozedural“, da vor allem Aktionen beschrieben werden. Das Konzept der imaginären Wanderung ist für die Inhaltsbestimmung das strukturierende Element. Im Gegensatz dazu werden Wegbeschreibungen für Einwohner „deklarativ“ genannt, da sie sich nicht auf Aktionen, sondern auf globale räumliche Sachverhalte beziehen. Demzufolge wird das strukturierende Element der imaginären Wanderung zugunsten einer effizienteren Informationsauswahl weitgehend aufgelöst. Besonders auffällig wird diese Dissonanz bei der typspezifischen Beschreibung von Wegen bei ihrer Analyse bezüglich Vollständigkeit der verbalisierten Information. Im Touristenfall werden nahezu immer alle Wege erwähnt, wohingegen für Einwohner teilweise große Wegabschnitte nicht verbalisiert werden. Offenbar gehen die Beschreibenden davon aus, daß diese Abschnitte eindeutig sind. In diesem Sinne vermuten Höök und Karlgren, daß deklarative Beschreibungen vor allem dazu dienen, Alternativen auszuschließen, weswegen die Wege weggelassen werden, die in der Nähe des Starts bzw. in der Nähe des Ziels liegen. Unterstützt wird eine solche Annahme dadurch, daß prozedurale Beschreibungen häufig auf Landmarken zurückgreifen. Bei der Auswahl des Weges stellen sie heraus, daß für Touristen Wege mit möglichst wenigen Abbiegevorgängen bevorzugt werden.

Diese stark heuristische Basis verwenden Höök und Karlgren, um einen Rahmen für ein Berechnungsmodell vorzustellen. Kernstück des Systems ist ein sogenanntes *route chunking* Modul, welches Wegabschnitte und Verzweigungspunkte entfernt und Wegabschnitte zusammenfaßt. Die Ausgabe dieses Moduls wird in einen Textgenerierungsprozeß eingegeben.

### 2.7.3 Das formale Modell von Hoepfner et al.

In allen bisher vorgestellten Arbeiten wird von einer strikten Zweiteilung in *Wegsuche* und *Sprachgenerierung* ausgegangen. Hoepfner und Mitarbeiter führen theoretische Überlegungen ins Feld, daß diese beiden Prozesse nicht rein sequentiell, sondern miteinander verzahnt agieren (vgl. [HCR90]). Sie geben Beispiele an, daß ein rein sequentielles Vorgehen, so wie es prinzipiell von Wahlster (vgl. [Wah82]) und Goerz (vgl. [Gö89]) beschrieben wird, beim Diskurstyp der vollständigen Wegbeschreibung nicht ausreicht (vgl. [HCR90]). Hierzu geben sie Beispiele an, in denen deutlich wird, inwieweit sogenannte *Interdependenzen* zwischen beiden Phasen zu berücksichtigen sind. Motiviert wird dieser Ansatz durch die Qualität einer Wegbeschreibung in Bezug auf *Merkbarkeit* und *Verifizierbarkeit*. Hoepfner et al. gehen davon aus, daß der Beschreibende berücksichtigt, daß ein Rezipient sich Landmarken und Aktionen besser merken kann, wenn die Landmarken in ihrem Typ variieren. Die Arbeiten von Hoepfner et al. sind erste Überlegungen im Hinblick auf eine Modellierung eines Generierungssystems von vollständigen Wegbeschreibungen und bedürfen empirischer Untersuchungen und einer Implementation.

---

<sup>64</sup>Vergleichbar zum Ansatz von Timpf et al. (vgl. [TVPE92])

### 2.7.4 Das formale Modell von Couclelis

Couclelis (vgl. [Cou95]) stellt den Rahmen eines Wegbeschreibungssystems vor, in dem die Konzepte eines mentalen Modells (vgl. [JL83], ähnlich zu Lakoffs “idealized cognitive models” (vgl. [Lak87]) und Fauconniers “mental spaces” (vgl. [Fau85])) mit dem Konzept der *Kognitiven Karte* (vgl. Kapitel 2.3.7) verbunden werden. Sie nimmt an, daß zur Generierung von Wegbeschreibungen das in mentalen Modellen repräsentierte spezifische und generelle Wissen über Umgebungen mit intentionalen Zuständen und Verhaltensformen integriert ist. Kognitive Karten, so postuliert Couclelis, sind ebenso wie Wegbeschreibungen in mentalen Modellen der Wegbeschreibungssituation verankert. Dies widerspricht der allgemeinen Annahme, daß Kognitive Karten die Grundlage von raumsprachlichen Beschreibungen sind. Mentale Modelle selbst sind wiederum in primitiven, kognitiven Strukturen, insbesondere “image-schemas” (vgl. [Tal83], [Lak87]) und “basic-level” Kategorien (vgl. [Ros73]), verankert und repräsentieren darüber die gesamte Wegbeschreibungssituation.

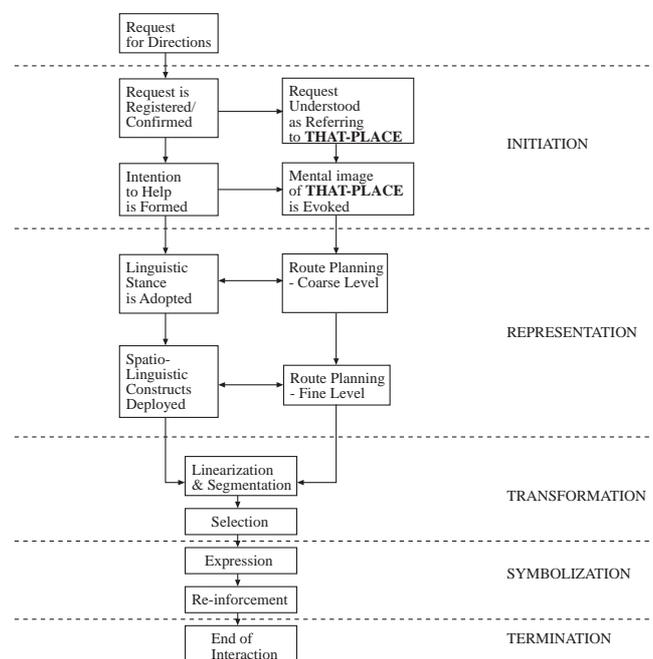


Abbildung 2.15: Die Stufen zur Generierung von Wegbeschreibungen (nach [Cou95])

Die Wegbeschreibung selbst wird durch ein *Komplexschema* gesteuert (siehe Abbildung 2.15). Das Modell zur Generierung einer vollständigen Wegbeschreibung besteht nach Couclelis aus den fünf Stufen der *Initialisierung*, *Repräsentation*, *Transformation*, *Symbolisierung* und *Termination*. Zuerst wird die Lokation des Ziels bestimmt. Das gesamte Wegbeschreibungsschema rekurriert auf die Aktivierung und Integration von schematischem Wissen. Wie von Klein (vgl. [Kle79]) vorgeschlagen, wird ein Pfad durch einen zweistufigen Prozeß von einem Primärplan (coarse level) in einen Sekundärplan (fine level) transformiert, bei dem eine allgemeine räumliche Repräsentation des Pfades basierend auf dem Überblickswissen des Sprechers in eine detaillierte Repräsentation umgesetzt wird. Die Linearisierung und Segmentierung der Beschreibung erfolgt über saliente Landmarken und Verzweigungspunkte (vgl. [AK85]).

Das Modell expliziert einen Rahmen für die Generierung von Wegbeschreibungen, welcher auf Überlegungen aus der Perspektive des “experiential realism” beruhen, der in Lakoff seinen prominentesten Vertreter hat. Mentale Modelle sind zwar abstrakter definiert als Kognitive Karten, übernehmen aber zum Großteil deren Aufgaben. Couclelis versucht Kognitive Karten nicht als

eigenständige, zwischen Sprache und Raum vermittelnde Repräsentationsformate, sondern neben raumbezogenen Diskursen als ein mögliches Explikationsformat aufzufassen. Es bleibt offen, in welcher Form räumliches Wissen in solcher Art mentalen Modellen repräsentiert wird und ob damit Phänomene der Kognitiven Karten (z.B. Verzerrungen) erklärt werden können.

### 2.7.5 Das Modell von Rau und Schweitzer

Ein elaboriertes Modell zur Generierung von vollständigen Wegbeschreibungen basierend auf einem schema- bzw. scriptbasierten Ansatz ist von Rau und Schweitzer vorgestellt worden (vgl. [RS87]). Entsprechend der Kleinschen Einteilung in Primär- und Sekundärpläne werden darin Transformationschritte angegeben, wie man von einer vollständigen Repräsentation von Wegabschnitten, über Zwischenschritte zu einer Beschreibung von Aktionen gelangt, die wiederum als Eingabe für einen Sprachgenerierungsmodule dient (vgl. Abbildung 2.16). Als Grundlage wird eine vollständige, geometrische Beschreibung von Landmarken und Straßenelementen und assoziierte Landmarkenmerkmale verwendet. Aus diesem Modell wird mittels eines nicht spezifizierten Wegfindungsalgorithmus eine Liste von Wegabschnitten festgelegt. Jedem einzelnen Abschnitt wird ein sogenanntes *Zustandsscript* zugeordnet. In diesem sind die Landmarken, insbesondere deren Lokation und Auffälligkeit entlang des Wegabschnitts vermerkt. Weiterhin ist der Endknoten des Wegabschnitts eingetragen sowie die relative Richtung zwischen Anfangs- und Endpunkt. Durch den nachfolgenden *Routengenerierungsalgorithmus* werden solche Abschnitte zusammenfaßt, zwischen denen ein Verzweigungspunkt oder ein großer Platz liegt. Durch mehrmalige Evaluation der *Regelanwendungsprozedur* erhält man eine Liste von sogenannten *Stepisoden*. Daraus wird durch einfache Transformationen ein *Routenscript* erzeugt. Rau und Schweitzer bezeichnen das Routenscript als Zustandsbeschreibung, die den Handlungsanweisungscharakter von Wegbeschreibungen noch nicht widerspiegelt. Um dies zu erreichen, wird das Routenscript ebenfalls durch Anwendung von Regeln in ein *Aktionsscript* umgesetzt, wobei die segmentierte Struktur des Routenscripts erhalten bleibt. Eine Einheit eines Aktionsscripts, besteht aus drei Teilen: einem *KS-*, einem *CD-MACRO* und einem *CD-MICRO* Teil. Der *KS-*Teil ist der Bedingungsteil eines Aktionsscripts. Basierend auf den CD-Graphen von Schank und Abelson (vgl. [SA77]) besteht der zweite Teil eines Aktionsscripts aus einem *CD-MICRO* und einem *CD-MACRO* Teil. Rau und Schweitzer nutzen für den *CD-MACRO* Teil ausschließlich die primitiven Aktionen *PTRANS* zur Repräsentation der Fortbewegung von einem Anfangs- zu einem Zielort und *ATTEND* als Repräsentation eines Wahrnehmungsaktes. Die *CD-MICRO* Repräsentation wird nur als Zusatzinformation im Bedarfsfalle eingeführt. In diesem Teil werden die im *CD-MACRO* Teil eingeführten Aktionen genauer spezifiziert. Letztendlich erhält man eine Liste von Aktionsscripts, die anschließend von einer semantischen Auswertungsprozedur evaluiert wird, bevor eine Texterzeugungskomponente die komplette Wegbeschreibung erstellt (vgl. [RS87]).

Nachteilig ist in diesem Modell, daß die Grundlage eine geometrische Repräsentation ist, in der alle Objekte und Straßen mit all ihren Merkmalen explizit repräsentiert sind. Da vollständige Wegbeschreibungen sich zumeist auf Erfahrungen über die jeweilige Umgebung stützen, stellt sich die generelle Frage nach der Vergleichbarkeit von kognitiven Karten und einem geometrischen Modell. Landmarken und deren räumliche Beziehungen zu Straßenabschnitten sind bereits im geometrischen Modell repräsentiert. Nur für die Bestimmung der Relation „zwischen“ wird eine einfache Prozedur angegeben (vgl. [RS87]). Neben den Lagebeziehungen von Landmarken enthält das Modell auch Information über fest kodierte Auffälligkeiten von Landmarken. Da eine Landmarke selten von allen Blickrichtungen eine homogene Verteilung der visuellen Merkmale besitzt, ist dieser Ansatz bezüglich der Verarbeitung von Merkmalen als starke Vereinfachung anzusehen. Die Eingabe besteht aus synthetischen Daten, die vorgegeben werden. Obwohl in den Beispielen die lokaldeiktischen Referenzen „hier“ und „dort“ verwendet werden, ist nicht beschrieben, wie solche Beschreibungen bestimmt werden. Das Gleiche gilt für tem-

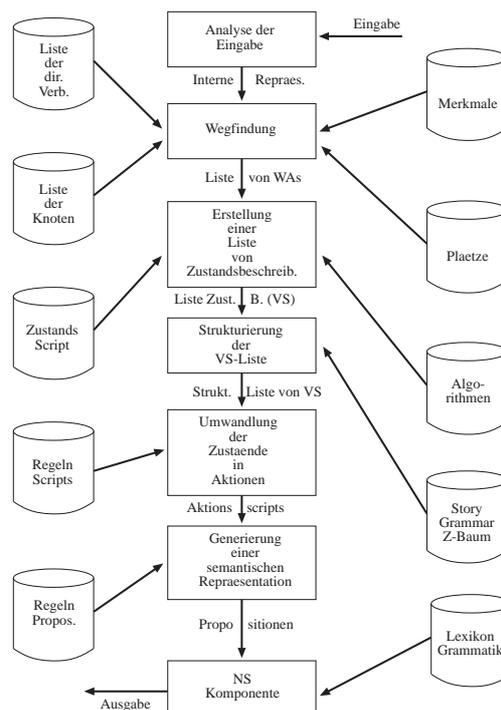


Abbildung 2.16: Systemarchitektur nach Rau und Schweitzer (nach. [RS87])

porale Bezüge wie „jetzt“. Obschon dem vorgeschlagenen prozeduralen Modell eine komplexe Architektur zugrunde liegt, kann keine Aussage zum Berechnungsaspekt gemacht werden, da das Modell nicht implementiert ist. Der wesentliche Aspekt des Modells von Rau und Schweitzer ist die Zusammenfassung von Wegabschnitten in Routenabschnitten. Obwohl eine Reihe von Transformationsschritten auf der Ausgangsliste ausgeführt werden, sind die Aktionscripte mit den Wegabschnitten inhaltlich und strukturell nahezu identisch.

### 2.7.6 Das Berechnungsmodell KLEIST

Ziel des Modells KLEIST ist es, mit besonderem Schwerpunkt auf der Erhaltung der inhaltlichen Kohärenz, solche kognitiven Prozesse im Kontext von Wegbeschreibungen zu modellieren, die beim Übergang des Primär- zum Sekundärplans ablaufen (vgl. [MMP<sup>+</sup>88], [RR90]). In empirischen Studien haben Meier et al. untersucht, in welcher Weise Probanden vollständige Wegbeschreibungen generieren. Als Eingabematerial ist eine textuelle Liste von Pfadabschnitten mit assoziierten Ausschnitten der Modellstadtkarte verwendet worden. Dadurch wurde den Probanden Information ähnlich eines Primärplans (vgl. [Kle79]) vorgegeben. Die generierten Beschreibungen einzelner Pfadabschnitte (Minimaltexte), so postulieren Meier et al., sind äquivalent zu Teilen eines Sekundärplans. Minimaltexte enthalten ausschließlich *Landmarken*, *Handlungsanweisungen* und *Richtungsangaben*. Abschließend wurden die Probanden aufgefordert, eine vollständige Wegbeschreibung aus den Einzelbeschreibungen abzuleiten. Minimaltexte lassen sich, entsprechend der Dreiteilung einer vollständigen Wegbeschreibung (vgl. [Kle79], [WR82]), der Anfangsetappe, der Zwischenstrecke oder der Zielbeschreibung zuordnen. Formal werden Minimaltexte durch schematische Teilbeschreibungen repräsentiert, die *Aktionen* und *Richtungen* mit Pfadabschnitten integrieren. In Teilbeschreibungen können mehrere Pfadabschnitte zusammengefaßt werden. Meier et al. gehen davon aus, daß jeder Minimaltext eine Angabe der Richtung, der Handlung und des Pfadabschnitts enthält.

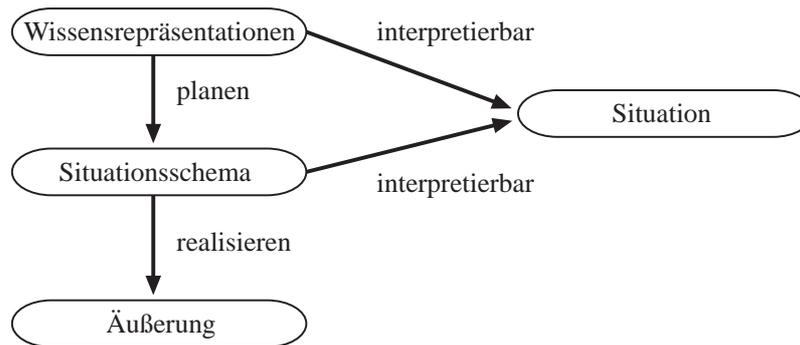


Abbildung 2.17: Repräsentationsstufen in KLEIST (nach [RR90])

Das Berechnungsmodell startet mit einer Liste von Kanten und Knoten, die Straßenabschnitte und Kreuzungen repräsentieren. Kreuzungen sind den Landmarken fest zugeordnet. Mittels dieser Information wird ein Teilbeschreibungstyp ausgewählt, der schrittweise verfeinert wird. Die Elemente einer Teilbeschreibung aktivieren Konzepte einer Konzepthierarchie (CheOPS), in der einerseits semantisches und syntaktisches und andererseits Wissen über die zugrundeliegende Modellumgebung abgelegt ist. Die Generierung einer Wegbeschreibung erfolgt durch passende Auswahl von Einheiten aus der Konzepthierarchie und der Kombination der annotierten Beschreibungen mittels Unifikation (siehe Abbildung 2.17). Das Ergebnis dieses Prozesses sind *Situationsschemata* (vgl. [FHLvB87]), die semantische und syntaktische Information einer Teilbeschreibung repräsentieren. Mittels Diskursregeln werden diese Situationsschemata durch Elliptifizierung, Fokussierung und Berücksichtigung von Regularitäten modifiziert. Abschließend werden durch einen LFG-Grammatikalischen Ansatz Oberflächenstrukturen generiert.

Das Ziel von KLEIST ist es, sprachproduktionsorientierte und nicht raumkognitive Prozesse zu untersuchen. In KLEIST werden eine Vielzahl von komplexen Problemen der Sprachverarbeitung analysiert. Die verwendete raumbezogene Information entspricht jedoch nur vage raumkognitiven Repräsentationen einer Kognitiven Karte. Beispielsweise werden räumliche Relationen statisch bezüglich einem egozentrischen Referenzsystems eines imaginären Wanderers zugeordnet. Weiterhin wird keine Operationalisierung zur Bestimmung räumlichen Wissens angegeben.

### 2.7.7 Das Berechnungsmodell CITYGUIDE

CITYGUIDE ist eines der ersten vollständig implementierten Modelle, welches eine natürlichsprachliche Beschreibung eines Weges ermöglicht (vgl. [Mül88], [Wah91]). Wie vergleichbare Ansätze (vgl. [HCR90], [Car91]), basiert dieses Modell auf einer netzartigen Repräsentation der räumlichen Umgebung.

CITYGUIDE ist eine Erweiterung des Systems CITYTOUR (vgl. [AHR89]) und steht somit in engem Zusammenhang zur Untersuchung, wie ein natürlichsprachlicher Zugang zu visuellen Daten erzielt werden kann (für einen Überblick siehe [Her96]). Die Aufgabe des Modells CITYGUIDE ist es, vollständige Wegbeschreibungen zu generieren. Das Modell unterteilt sich in eine *Wegfindungs-* und eine *Wegbeschreibungskomponente*. Der Wegfindung liegt ein gerichteter Graph zugrunde, welcher gleichzeitig den relevanten Kartenausschnitt und die Umgebung repräsentiert. Ein Knoten des Graphs repräsentiert den Anfangs- oder Endpunkt einer Kante, den Mittelpunkt einer Kreuzung oder eines Platzes oder einen Verbindungspunkt zu einer Landmarke. Eine Kante verbindet jeweils zwei Knoten und repräsentiert somit einen Wegabschnitt. Besonderer Fokus liegt in CITYGUIDE auf der Formalisierung von Plätzen, wobei Kreuzungen als eine Spezialform gelten. Ein Platz ist durch einen Knoten modelliert, welcher durch eine

Kante mit allen Knoten verbunden ist, die Endpunkte von in den Platz einmündende Straßen repräsentieren. Jeder dieser Endknoten ist untereinander verbunden.<sup>65</sup>

Neben dieser Art von Plätzen werden auch solche verwendet, die Landmarken repräsentieren. Ist eine Landmarke in der Nähe eines Anfangs- oder Endknotens eines Wegabschnitts, wird der Landmarkenknoten mit diesem Wegeknoten durch die Einführung einer neuen Kante verbunden. Liegt eine Landmarke auf einem Platz, wird der Landmarkenknoten mit dem Mittelpunktknoten und allen Zugangsknoten des Platzes verbunden. Wie bei KLEIST enthält die Repräsentation einer Landmarke, ob sie ein auffälliges Merkmal besitzt. Landmarken dienen in CITYGUIDE als Bezugspunkte für Pfade und Richtungsänderungen. Um Landmarken besser beschreiben zu können, werden Merkmale berücksichtigt, die eine „gewisse“ Auffälligkeit besitzen. Wie jedoch die Auffälligkeit spezifiziert wird, bleibt offen. Der kürzeste Weg zwischen einem Start- und Zielknoten wird durch den Dijkstra Suchalgorithmus bestimmt. Entsprechend dem Modell von Klein faßt das Wegbeschreibungsmodul solche Wegabschnitte zu Routenabschnitten zusammen, die demselben *Straßennamen* zugeordnet sind.<sup>66</sup> Die Generierung der Wegbeschreibung ist schemabasiert. Die einzelnen Übergänge von einem Routenabschnitt zum nächsten erfolgen durch Nennung der Richtung oder des folgenden Straßen- oder Platznamens und, falls vorhanden, der Beschreibung eines auffälligen Merkmals. Befindet sich an der Übergangsstelle keine Landmarke, so wird per Numerierung angegeben, welche Abzweigung genommen werden soll. Beschreibungen einzelner Routenabschnitte sind mittels „und“ verbunden. Liegt eine Landmarke eines Routenabschnitts jedoch nicht in der Nähe dessen Anfangs- oder des Endknoten, wird als Beschreibung der Landmarke die Konstruktion „an ... vorbei“ oder „an ... entlang“ verwendet.

Die klare Aufteilung in einen Primär- und Sekundärplan, so wie sie von Klein vorgeschlagen wurde, wird in CITYGUIDE nur rudimentär durch Zusammenfassung von Pfadabschnitten zu Routen über die Indizierung durch den Straßennamen berücksichtigt. Routenabschnitte werden durch einfache Konkatenation miteinander verbunden, was dazu führt, daß längere Beschreibungen sehr künstlich und monoton klingen. Unverständlich werden die Beschreibungen sogar dann, wenn Rezipienten, Straßen- oder Landmarkennamen genannt werden. Ebenso nicht berücksichtigt werden lokaldeiktische und temporale Ausdrücke. Da CITYGUIDE sich ausschließlich auf Daten bezieht, die dem Wegenetz entstammen, wird die Beschreibung, metaphorisch gesehen, aus einer Vogelperspektive beschrieben. Routenwissen wird nicht verwendet, was auch dadurch zum Ausdruck kommt, daß Merkmale von Landmarken nur einen konstanten Auffälligkeitswert besitzen.

### 2.7.8 Das Berechnungsmodell von Carstensen

Carstensen (vgl. [Car91]) stellt einen theoretischen Rahmen für die Generierung von Wegbeschreibungen vor (vgl. Abbildung 2.18). In dieser Arbeit wird ebenfalls von Kleins Unterteilung in Wege und Routen ausgegangen (vgl. [Kle79]). Die Architektur des Wegbeschreibungsmodells basiert auf Levelts Unterteilung in Makro- und Mikro-Planung (vgl. [Lev89]). Die Bestimmung einer Route erfolgt in einem *Routenfindungsmodul*. Als Vermittlung zwischen der räumlichen und der sprachlichen Ebene wird eine konzeptuelle Ebene eingeführt (vgl. [BL87]). Dieses Modul und ein solches zur Bestimmung von Nebeninformation sind Bestandteil der Makro-Planung. Über Elemente der Mikroplanung werden keine spezifischen Aussagen gemacht. Der Fokus dieses Modells liegt auf der Interaktion von Routenfindung, Routennetz und relevanten konzeptuellen Kategorien.

---

<sup>65</sup>Es bleibt unklar, weshalb eine solche komplexe Repräsentation eingeführt wurde. Die Zusammenfassung aller Einmündungsknoten in einen Platz erfüllt offenbar ebenfalls die in CITYGUIDE betrachteten Erfordernisse.

<sup>66</sup>Eine derartige Strukturierung von pfadbezogenem Wissen ist ad hoc. Empirische Untersuchungen zeigen, daß eine Strukturierung von Pfaden nicht von symbolischen Markierungen abhängt (vgl. [AK85]).

Zentral ist der Begriff einer *konzeptuellen Kategorie*, die eine nicht-sprachliche, konzeptuelle Repräsentation von Gestalt- und Lageeigenschaften räumlicher Objekte beschreibt (vgl. [LCS90]). Eine sogenannte *Ereigniskategorie* dient der Auswahl von Lexemen und bildet eine Schnittstelle zwischen Repräsentationen auf räumlicher und sprachlicher Ebene. Ereigniskategorien sind schematische, sprachnahe Repräsentationen von Ereignissen, welche intern in verschiedene Auflösungsebenen unterteilt sind. Eine Routenkategorie ist ein Spezialfall einer Ereigniskategorie. Da Routenabschnitte nicht statisch, sondern sowohl zeitlich, als auch räumlich ausgedehnt sind, postuliert Carstensen sogenannte *Routenkategorien* zur konzeptuellen Zusammenfassung von Einzelsituationen nach dem Prinzip der *Ähnlichkeit*. Dabei besteht eine enge Verwandtschaft zwischen einer Ereigniskategorie und einer Folge von Aktionen (actions) und Sichten (views) im Sinne von Kuipers (vgl. [Kui78]). Zentral für die Routenfindung ist ein Wegfindungsprozeß, der durch einen bidirektionalen A\*-Algorithmus modelliert wird. Wird ein neuer Wegabschnitt bestimmt, so entscheidet der Routenkategorisierungsprozeß, ob „eine gerade bearbeitete Kante des Suchgraphen als zur aktuellen Routenkategorie gehörig einzuordnen ist oder es die Konstruktion einer neuen Routenkategorie erfordert“ (vgl. [Car91, S. 97]). Offen bleibt zum einen, wie diese Überprüfung erfolgt und zum anderen, wie eine neue Kategorie konstruiert werden kann. Das Ergebnis dieser planbasierten Routenfindung ist letztendlich eine Liste von Ereigniskategorien (ähnlich zu [RS87]). Basierend auf empirischen Untersuchungen (vgl. [Cha82], [Pai70]) und anderen Berechnungsmodellen (vgl. [LZ89]), wird eine *erfahrungsbasierte Routenfindung* angedeutet. Ist eine Route gefunden worden, so wird sie in einem Routennetz integriert. Ein Routennetz unterteilt sich in ein Basis- und ein Sekundärnetz (entsprechend [LZ89]). Im Basisnetz sind alle Routen enthalten, die häufig und im Sekundärnetz die, die weniger häufig verwendet werden. Die Wegsuche präferiert Routen des Routennetzes und dabei solche des Basisnetzes.

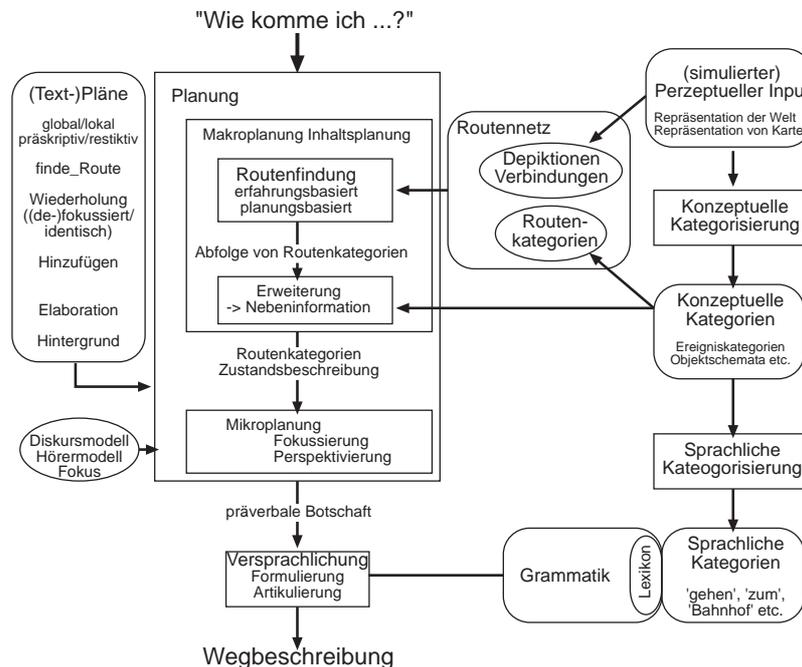


Abbildung 2.18: Das Modell der Generierung von Wegbeschreibungen (nach [Car91])

Das Modell ist eine Anwendung der *2-Stufen-Semantik* (vgl. [BL87]) auf den Bereich der Wegbeschreibungen. Hierzu wird eine konzeptuelle Ebene eingeführt. Konzeptuelle Strukturen besitzen einerseits eine vermittelnde Rolle zwischen der perzeptuellen und andererseits der motorischen Ebene und der sprachlichen Ebene. Jedoch bleibt Carstensen besonders für die Verbindung

zur perzeptuellen Ebene jedwede Evidenz schuldig. Ebenso offen bleibt die Frage, wie neue Kategorien konstruiert werden können, obschon dies in dieser Arbeit als besondere Neuerung herausgestellt wird. Somit scheint der vorgestellte Ansatz nicht über den schemabasierten Ansatz von Ruhrberg et al. (vgl. [RR90]) hinauszugehen. Gänzlich offen bleibt die Frage nach der Bestimmung von Routenkategorien. Hierbei bezieht sich Carstensen auf die *Ähnlichkeit* einer Situation. Wenn eine *Sicht* ihrer direkt vorausgegangenen *Sicht* ähnlich ist, gilt sie als Bestandteil der aktuellen Routenkategorie. Es wird aber kein Konzept für *Ähnlichkeit* angegeben. Letztendlich läßt sich aus der vorliegenden Arbeit nur schließen, daß Ereigniskonzepte im wesentlichen schematische Repräsentationen sind, die dazu dienen, lexikalische Einträge effizient und speziell für Wegbeschreibungen auszuwählen. Da das Modell nur teilweise implementiert ist, nimmt die Arbeit eine Zwischenstellung zwischen einem rein theoretischen und einem implementierten Modell ein.

### 2.7.9 Das *Back Seat Driver* Modell

Als Hilfesysteme für Autofahrer sind Wegbeschreibungssysteme in den vergangenen Jahren vielfach untersucht worden. Technische Lösungen sind jedoch durchweg nur an einer kostengünstigen Realisierung bei gleichzeitig minimalem technischen Aufwand interessiert. Hinter einer solchen Prämisse treten Überlegungen zur Benutzerfreundlichkeit zurück.<sup>67</sup>

Ein Mittelweg zwischen der Betrachtung technischer Möglichkeiten und Benutzerfreundlichkeit ist in der Arbeit von Davis versucht worden (vgl. [Dav89]). Das von Davis vorgestellte "Back Seat Driver" Modell beschreibt technische und benutzerorientierte Aspekte zur Generierung eines automatischen inkrementellen Wegbeschreibungssystem zur Fahrernavigation. Die Beschreibungen des Systems basieren auf einer Kartenrepräsentation. Davis identifiziert das Problem, daß für inkrementelle Wegbeschreibungen neben dem Inhalt auch der Zeitpunkt einer Beschreibung determiniert werden muß. Davis verwendet zur Beschreibung einer Aktion an einem Verzweigungspunkt neun verschiedene Erkennungsautomaten, die entsprechend der Konfiguration der Pfadabschnitte ausgewählt werden. Der Ort einer Beschreibung berechnet sich, entsprechend physikalischer Gesetze, aus der Multiplikation der Geschwindigkeit mit der Zeit, die benötigt wird, die Beschreibung zu präsentieren, addiert mit der Reaktionszeit des Rezipienten (2 Sekunden).

Das Back Seat Driver Modell ist eines der ersten Modelle, welches die automatische Generierung von inkrementellen Wegbeschreibungen genauer untersucht hat. Der Wert dieser Arbeit ist nicht genau zu bestimmen, da Davis über eine deskriptive Darstellung des Modells nicht hinauskommt. Weder stellt er vor, wie die Erkennungsautomaten für Aktionen ausgewählt, noch wie Beschreibungen generiert werden. Letztendlich ist die Darstellung des Modells eine reine Auflistung von Elementen, die ein inkrementelles Wegbeschreibungssystem enthalten müßte. Ob die Elemente im Back Seat Driver integriert sind, ist nicht zu entscheiden. Das Modell entspringt einer rein technischen Betrachtungsweise. Untersucht wird inwieweit Sprache im Kontext von Wegbeschreibungen an Stelle von Graphiken verwendet werden kann. Bedingt durch den Einsatz in einem Automobil sind in diesem Modell erste einfache Funktionalitäten integriert worden, die den Inhalt der Beschreibung an zeitlichen Beschränkungen während der Fahrt anpassen. In diesem Sinne ist das Back Seat Driver Modell eines der ersten Ansätze, um Wegbeschreibungen inkrementell und ressourcenadaptiv zu generieren.

---

<sup>67</sup>Der TravelPilot von Blaupunkt ist nur eins von vielen Beispielen.

WB-Systeme	Habel	Höök	Hoepfner	Couclelis	Rau	KLEIST	CITYGUIDE	Carstensen	Davis	MOSES <sup>a</sup>
Vollständige Wegbeschreibungen	•	•	•	•	•	•	•	•		
Inkrementelle Wegbeschreibung						○			•	•
Primär/Sekundärplan	•		•	•	•	•	•	•		
Geometrische Umgebung	•	•	•		•	○	•	•	•	•
“dual coding” Repräsentation	•			•						
Vollständiges räumliches Wissen	•	•	•		•	•	•	•	•	
Limitiertes räumliches Wissen				•						•
Egozentrisches Referenzsystem	•	•	○	•	•	•	•	•	•	•
Allozentrisches Referenzsystem				•						•
Imaginäre Wanderung <sup>b</sup>	•		•	•	•	•	•		•	
Berechnung räumlicher Relationen	•		•		○	•	•	•		•
Gradierte räumliche Relationen	○						•			•
Verschiedene Wegetypen		•								•
Verschiedene Rezipiententypen		•							•	
Landmarken	•	•	•	•	•	•	•	•	•	•
Feste visuelle Salienz	•	•	•		•		•			
Berechnete visuelle Salienz				•						•
Plätze	•		•				•			•
Räumliche Nähe						○	○	•		•
Schemabasierte Sprachgenerierung	•		○		•	•	•	•	•	•
Empirische Studien	•	•		•		•		•	•	•
Implementiert	○	•	○			•	•	○	○	•

<sup>a</sup>Name des implementierten Agenten des in Kapitel 4 beschriebenen funktionalen Modells

<sup>b</sup>Imaginäre Wanderung

Tabelle 2.4: Vergleich der formalen Modelle zu Wegbeschreibungen: • indiziert Modellierung entsprechender Funktionalitäten; ○ indiziert partielle Modellierung

### 2.7.10 Zusammenfassung der Wegbeschreibungsmodelle

Die hier vorgestellten Modelle untersuchen einerseits Fragestellungen im Kontext von Wegbeschreibungen und andererseits allgemeine Mechanismen zur Verarbeitung räumlichen Wissens (vgl. Tabelle 2.4). Bis auf das *Back Seat Driver* Modell von Davis, untersuchen alle anderen Arbeiten der ersten Kategorie vollständige Wegbeschreibungen. Dies ist einerseits dadurch verständlich, daß dabei nur geringe dynamische Aspekte zu betrachten sind, was nicht zuletzt die empirische Datenerhebung vereinfacht. Andererseits basieren die meisten genannten Modelle zu Wegbeschreibungen auf nur ungenau spezifizierten Repräsentationen räumlichen Wissens. Hiervon ist Habels Ansatz ausgenommen, dessen Repräsentationsformalismus der Referentiellen Netze äußerst dediziert ist, jedoch ein Wegbeschreibungsmodell nur skizziert. Die meisten Arbeiten im Kontext der Wegbeschreibungen fokussieren vor allem auf sprachliche Aspekte, so daß Fragen der Raumrepräsentation in den Hintergrund treten.

Vollständige Wegbeschreibungen basieren auf Langzeitspeicherungen räumlichen Wissens. In den vorgestellten formalen Modellen wird keine Aussage darüber gemacht, woher dieses Wissen kommt und wie es aufgebaut wird. Durchgängig sind die Repräsentationen so gewählt, daß sie bestimmte Aufgaben, insbesondere sprachliche, gut unterstützen. Aus Ingenieursicht ist dagegen nichts einzuwenden. Besteht aber, wie bei allen diesen Modellen, der Anspruch einer mehr oder weniger direkten Simulation bzw. Nachbildung kognitiver Leistungen, ist die Frage nach dem Ursprung räumlichen Wissen fundamental.

Vor diesem Hintergrund bieten inkrementelle Wegbeschreibungen einen Ansatzpunkt, zu untersuchen, wie unmittelbar perzipierte Information verwendet wird, um räumliches Wissen aufzubauen und sprachlich zu verwenden. Zusätzlich sind aber auch zeitliche Aspekte zu betrachten, die durch die Bewegung des Sprechers und der Umgebung selbst induziert werden.

## 2.8 Zusammenfassung

In diesem Kapitel sind Elemente indiziert worden, die in einer vollständigen Theorie der Raumkognition enthalten sein müssen. Bisher ist nur schemenhaft zu erkennen, wie eine solche Theorie aussehen kann. In der empirischen Literatur herrscht weitgehend Einigkeit darüber, daß es amodale, eigenständige Repräsentationen räumlichen Wissen gibt, die aus unterschiedlichen Perzeptionsfunktionen abgeleitet werden. In der sprachwissenschaftlichen Literatur herrscht diese Einigkeit nicht in einer solch eindeutigen Weise. In einflußreichen Ansätzen (z.B. [Jac87], [Her86]) wird implizit oder explizit die Auffassung vertreten, daß räumliches Wissen Teil des grammatikalisch-lexikalischen Wissens ist. Im Bereich der Künstlichen Intelligenz gibt es Arbeiten, die die eine oder andere Perspektive einnehmen. Letztendlich kann diese Frage bisher nicht mit endgültiger Sicherheit beantwortet werden, was nicht zuletzt auf der Komplexität räumlichen Wissens beruht.

Weitgehend einig sind sich die vorgestellten Arbeiten in der konzeptuellen Zerlegung räumlichen Wissens. Grundlegend ist das Konzept des Objektes und der Lokation. Weiterhin enthält räumliches Wissen inhärent das Konzept einer räumlichen Relation, die Objekte bzw. Lokationen zueinander in Beziehung setzt. Nicht entschieden ist, ob räumliche Relationen zwischen Ausprägungen der Objekte selbst oder den von ihnen eingenommenen Lokationen etabliert werden. Die Etablierung räumlichen Wissens bedarf jedoch grundlegend einer Perspektive. Im primären Fall ist die Perspektive durch den Betrachter/Sprecher selbst gegeben, jedoch besitzt der Mensch auch die Fähigkeit, sich in andere Objekte und Lokationen einer Umgebung imaginär hineinzuversetzen. Das Konzept, welches hierzu verwendet wird, ist das des Referenzsystems. Häufig wird ein Referenzsystem mit einem kartesischen Referenzsystem gleichgesetzt. Dies schränkt das Konzept stark ein und bezieht sich fast ausschließlich auf die Körperachsen des Menschen und

den sprachlichen Gebrauch räumlicher Relationen. Andere Referenzsysteme, wie beispielsweise polare, scheinen aber auch in bestimmten Situationen Verwendung zu finden. Bisher unbeachtet blieb in den Arbeiten zu Referenzsystemen die Fragestellung, wie und warum von einem Referenzsystem zu einem anderen gewechselt wird.

Mit diesen Konzepten Objekt, Lokation, räumliche Relation und Referenzsystem läßt sich räumliches Wissen analysieren. Orthogonal zu diesen Konzepten ist die qualitative und quantitative Strukturierung des Raums. Bedingt durch mathematische Formalismen sind quantitative Formalisierungen der einzelnen Konzepte häufig untersucht worden. Assoziiert sind Metriken, bezüglich denen Distanzen und Winkel quantifiziert werden. Eng verbunden mit quantitativen Formalisierungen räumlicher Konzepte ist die Repräsentation durch geometrische Konzepte. Wie empirische Arbeiten zur Wahrnehmung zeigen, lassen sich durch geometrische Konzepte räumliche Vorstellungen in effizienter und effektiver Weise repräsentieren. Besonders im Hinblick auf visuelle Wahrnehmung repräsentieren geometrische Konzepte in geeigneter Weise das, was ein Akteur „sieht“. Dies ist Ursache dafür, daß in den formalen Ansätzen geometrische Repräsentationen als Grundlage verwendet werden.<sup>68</sup> Nicht formalisiert ist dabei durchgehend, wie eben eine solche geometrische Grundlage aufgebaut wird. Arbeiten, die eine Anbindung an perzeptuelle Eingabekanäle, wie beispielsweise visuelle Daten, besitzen, geben hierüber Aufschluß. In geometrischen Repräsentationen ist kein räumliches Wissen explizit repräsentiert. Erst die Etablierung von räumlichen Relationen bei Verwendung von Referenzsystemen erreicht diese Explikation.

Empirische Arbeiten bauen auf den genannten Konzepten auf und unterscheiden verschiedene Arten räumlichen Wissens, wie Landmarken-, Routen- und Konfigurationswissen. Die Entwicklung dieser Wissensarten bedarf Erfahrung über Umgebungen, weswegen gedächtnisspezifische Aspekte bei der Memorierung und der Erinnerung einen wesentlichen Einfluß besitzen.

---

<sup>68</sup>Ausgenommen ist hiervon das TOUR Modell.

## Kapitel 3

# Empirische Untersuchungen zu inkrementellen Wegbeschreibungen

Obschon Sprache starken individuellen Unterschieden unterliegt, ermöglicht sie selbst in restriktiven Situationen eine effiziente Vermittlung schwieriger Sachverhalte. Dies beruht vor allem auf der Konventionalisierung von Kommunikationshandlungen. Beschreibungen werden vom Rezipienten inferiert, welches wiederum dem Sprecher erlaubt, sich auf die Vermittlung der wesentlichen Information zu beschränken. Sprachliche Ausdrücke werden vielfach nur noch als Auslöser zum Verständnis komplexer Inhalte verwendet. Eine Anweisung „Da, rechts!“ vermittelt i.a. einem Rezipienten genügend Information, um sich in einer gegebenen Situation zu orientieren, Objekte zu extrahieren und motorische Aktionen auszuführen. Wie in zahlreichen Studien gezeigt worden ist, sind Wegbeschreibungen ein Beispiel für besonders stark schematisierte Kommunikationshandlungen (vgl. Kapitel 2.5.2). Basierend auf unterschiedlichen, inhaltlichen Kategorien sind in diesen Arbeiten semantische und syntaktische Strukturen von vollständigen Wegbeschreibungen klassifiziert worden.

In Bezug auf inkrementelle Wegbeschreibungen sind empirische Untersuchungen bisher nicht durchgeführt worden. Aus diesem Grund ist als Startpunkt im Rahmen dieser Arbeit durch eine empirische Untersuchung ein Sprachkorpus inkrementeller Wegbeschreibungen erstellt und analysiert worden. Da inkrementelle Wegbeschreibungen alltägliche kommunikative Handlungen sind, ist die Untersuchung in einer Realumgebung durchgeführt worden, was eine hohe ökologische Validität gewährleistet. Bedingt durch die geringe Kontrollierbarkeit von Situationseinflüssen in der ersten Untersuchung, ist in einer zweiten unter Laborbedingungen eine einzelne Verkehrssituation verwendet worden, um Auswirkungen kontrollierter Parametervariationen auf das Sprachverhalten zu überprüfen. Im weiteren wird nach einigen Vorbemerkungen die Felduntersuchung und anschließend die Laboruntersuchung beschrieben.

### 3.1 Vorbemerkungen

Die Kommunikationshandlung, die durch *inkrementelle Wegbeschreibungen* gegeben ist, wirft eine Reihe empirischer Fragen auf. Aus linguistischer Sicht sind vollständige Wegbeschreibungen häufig auf ihre grammatikalische Struktur untersucht worden (vgl. Kapitel 2.5.2). Solche Beschreibungen beruhen auf Situationen, die vom Sprecher auf Grund seiner Erfahrung erinnert werden. Die Aufgabe der Probanden besteht in solchen Untersuchungen darin, von memoriertem, raumbezogenem Wissen Beschreibungen abzuleiten. Bei der Generierung vollständiger Wegbeschreibungen sind in erheblichem Maße Phänomene zu berücksichtigen, die auf Inferenz-

schritten des Sprechers über den Rezipienten beruhen (z.B. [HK91], [Kle79]).<sup>1</sup> Da der Rezipient eine mentale Vorstellung alleine auf der Basis der Beschreibungen ableitet, sind diese vom Sprecher derart zu strukturieren, daß sie in ausreichendem Maße memoriert werden können und später verifizierbar sind (vgl. [HCR90]). Es vermischt sich in vollständigen Wegbeschreibungen somit das allgemeine Problem der Generierung von Raumbeschreibungen mit solchen, die sich auf Memorierungsleistungen beziehen. In vergleichbaren Arbeiten ist untersucht worden, welches raumbezogene Wissen Probanden bei einer Autofahrt durch eine fremde Umgebung aufbauen (z.B. [ALM64], [CS69]). Dabei zeigt sich eine große Übereinstimmung bei den memorierten Objekten (vgl. [CS69]), was die kommunikative Verwendbarkeit von solchen Objekten erhöht. Im Kontext der inkrementellen Wegbeschreibungen verifiziert der Rezipient die aus einer Beschreibung abgeleitete Bedeutung direkt in der perzipierbaren Umgebung, so daß der Aufwand der Gedächtnisleistungen erheblich verringert wird.

Maßgeblich für die Generierung inkrementeller Wegbeschreibungen ist die Eigenbewegung des Sprechers und des Rezipienten. Damit verknüpft sind Abschätzungen über Geschwindigkeiten und die Verwendung anderer wahrnehmbarer Eigenschaften einer Umgebung. Vor diesem Hintergrund lassen sich für inkrementelle Wegbeschreibungen (IWB) folgende Hypothesen aufstellen:

1. Inhaltsstrukturen von IWB sind aus einer geringen Menge von sprachlichen *Inhaltskategorien* aufgebaut.
2. Inhaltliche und sprachliche Strukturen von IWB sind an Zeitbeschränkungen der Situation angepaßt.
3. Die Referenzierbarkeit eines Objektes setzt eine ausreichend hohe visuelle Salienz voraus.
4. Objekte werden nur dann referenziert, wenn sie sich an geeigneten räumlichen Lokationen befinden.

Die *erste Hypothese* postuliert, daß sich inkrementelle Wegbeschreibungen aus kommunikativen Schemata aufbauen lassen, die auf einer geringen Menge von Inhaltskategorien bestehen (z.B. [Hab87], [RS87]). In vollständigen Wegbeschreibungen lassen sich entsprechend Lynchs Ontologie fünf Kategorien identifizieren (vgl. [Lyn60]): Landmarken, Pfade, Verzweigungspunkte, Grenzen und Gebiete. Besitzt der Sprecher bezüglich der zu beschreibenden Umgebung kein raumbezogenes Wissen, bezieht er sich in inkrementellen Wegbeschreibungen ausschließlich auf die ersten drei Kategorien (vgl. auch [Kle79]). Aus diesen Kategorien lassen sich Inhaltskategorien ableiten, auf denen inkrementelle Wegbeschreibungen aufgebaut sind. Bezüglich stadtähnlichen Umgebungen wird davon ausgegangen, daß Landmarken statisch<sup>2</sup> an einer Lokation verankert sind (z.B. [Lyn60], [ALM64] und vgl. Kapitel 2). Eine Beschreibung einer solchen Lokation wird durch eine *WO-K*<sup>3</sup> Inhaltskategorie repräsentiert. Eine Beispielbeschreibung lautet wie folgt: „Vor dem großen Haus auf der rechten Seite ...“ Eine Lokation kann aber auch alleine durch eine deiktische Referenz indiziert werden: „Hier ...“ Eine weitere Inhaltskategorie, *WAS-K*, wird durch die Beschreibung von *Pfaden* impliziert. Jedem Pfad sind verschiedene Bewegungsmodi zugeordnet, welche Einfluß auf das Verb besitzen (z.B. „laufen“, „gehen“ und „fahren“). Die wesentliche Bedeutung einer inkrementellen Wegbeschreibung basiert auf der einzuschlagenden Richtung.

---

<sup>1</sup>In einer diagonal hierzu angelegten Untersuchung haben Streeter und Kollegen ([SVW85]) die Verwendung verschiedener Informationsmodi als Fahrernavigationshilfe verglichen. Von einem Kassettenrekorder inkrementell gegebene Anweisungen sind durch die Versuchsleiter vorgegeben worden. Sie fanden heraus, daß verbale Beschreibungen effektiver vom Rezipienten verwendet werden können als eine Straßenkarte oder einer gleichzeitigen Verwendung einer verbalen Beschreibung und einer Straßenkarte.

<sup>2</sup>Dynamische Landmarken, wie beispielsweise der Haleysche Komet, werden selten verwendet.

<sup>3</sup>Die Erweiterung durch '-K' indiziert, daß es sich um eine Inhaltskategorie handelt.

Die Richtung am nächsten Verzweigungspunkt wird durch den Pfadabschnitt bis zum Verzweigungspunkt und dem daran anschließenden festgelegt. Diese Information über die Richtung wird durch die *WOHIN-K* Inhaltskategorie klassifiziert. *WOHIN-K* Kategorien sind eng mit *WAS-K* Kategorien verbunden, da sie im Deutschen häufig durch eine adverbiale Ergänzung ausgedrückt werden. Ein Beispiel hierfür ist die folgende Beschreibung: „Nach rechts abbiegen.“ Hierbei ist „abbiegen“ eine Instanziierung einer *WAS-K* und „nach rechts“ einer *WOHIN-K* Kategorie.

Im Gegensatz zu sprachlichen Untersuchungen vollständiger Wegbeschreibungen, sind bei der Generierung inkrementeller Wegbeschreibungen verstärkt zeitliche Referenzen und zeitliche Restriktionen zu berücksichtigen. Sprachliche Ausdrücke, die zeitbezogene Referenzen indizieren, werden unter die Inhaltskategorie *WANN-K* subsummiert. Ebenso wie raumbezogene Referenzen werden zeitliche vor allem zum Aufbau deiktischer Räume verwendet (z.B. „jetzt“ und „gleich“). Dabei werden zeitliche und räumliche Referenzen synonym verwendet (z.B. [Kle79]). Zeitbezogene Ausdrücke beschreiben lineare Sequenzen von Einzelaktionen (z.B. „dann“). Sie können aber auch den Zeitpunkt ankündigen, an dem eine Aktion auszuführen ist (z.B. „jetzt“). Diese vier Inhaltskategorien werden in der ersten empirischen Untersuchung bezüglich inkrementellen Wegbeschreibungen untersucht. Weiterhin wird überprüft, inwieweit diese Kategorien die Gesamtheit des dabei gewonnenen Korpus abdecken, bzw. welchen Stellenwert eine Restkategorie besitzt. Auf Grund der schematischen Konstruktionsweise von vollständigen Wegbeschreibungen wird überprüft, in welcher Weise Inhaltskategorien miteinander zu komplexen Beschreibungen integriert werden. Die *zweite Hypothese* betrifft die Frage, in welcher Weise sich Effekte zeitlicher Restriktionen in den Beschreibungen wiederfinden lassen, die vor allem durch die Eigenbewegung des Sprechers und des Rezipienten bewirkt werden. Zu vermuten ist eine Verkürzung der Beschreibungen, je näher beide dem nächsten Verzweigungspunkt sind. In der ersten empirischen Untersuchung wird analysiert, ob dies zu einer charakteristischen Verwendung von Inhaltsstrukturen führt.

Zusammenfassend lassen sich die vier Inhaltsstrukturen wie folgt charakterisieren:

1. *WOHIN-K*: indiziert die Richtung der auszuführenden Aktion
2. *WO-K*: referenziert über die Beschreibung eines Objektes oder eines Pfadelementes ein Gebiet
3. *WANN-K*: gibt eine qualitative<sup>4</sup> Beschreibung des Zeitpunktes zur Ausführung einer Aktion
4. *WAS-K*: referenziert die Art der Bewegung

Im Bereich der Forschung zur visuellen Wahrnehmung sind Aufmerksamkeits- und Fokussierungsprozesse in zahlreichen Studien untersucht worden. Der Einfluß solcher Aufmerksamkeitsprozesse auf Prozesse der sprachlichen Ebene ist hingegen bisher weniger untersucht worden (z.B. [GS86], [AK87]). Eine Fragestellung, die aus einer integrierten Betrachtungsweise von visueller Wahrnehmung und Sprache entsteht, ist, in welcher Weise visuell fokussierte Entitäten in sprachlichen Beschreibungen verwendet werden, aber auch, ob und wie sprachliche Prozesse Aufmerksamkeitsprozesse beeinflussen (vgl. [BJKZ85]). Solche Fragestellungen führten zu empirischen Untersuchungen (z.B. [Coh84], [CMP<sup>+</sup>95]) und formalen Modellierungen (z.B. [App85b], [RD92], [HH92], vgl. Kapitel 2.5). Eine der Fragen in diesem Kontext betrifft die Auswahl der Kriterien für visuell perzipierte Objekte. In der Aufmerksamkeitsforschung werden momentan

---

<sup>4</sup>Quantitative Beschreibungen sind ebenfalls möglich. Sie traten jedoch im verwendeten Korpus nicht auf. Durch den Einsatz von technischen Geräten im Alltag, wie beispielsweise GPS, ist es denkbar, daß quantitative Beschreibungen in Zukunft immer häufiger verwendet werden.

merkmalsbasierte Ansätze favorisiert, da gestaltpsychologisch-holistische Theorien Formalisierungen nur schwerlich ermöglichen. Basierend auf der “feature integration theory” von Anne Treisman (vgl. Kapitel 2) ist in einer zweiten empirischen Untersuchung analysiert worden, welche sprachlichen Auswirkungen Variationen von visuell wahrnehmbaren, physikalischen Merkmalen in einer idealisierten Verkehrssituation haben. Es wird untersucht, ob visuelle Merkmale eine sprachliche Referenzierung beeinflussen. Dabei wird postuliert, daß neben visuellen Merkmalen auch die räumliche Lokation eines Objekts beeinflußt, ob es als salient betrachtet wird. Es wird vermutet, daß ein Objekt erst dann referenziert wird, wenn es eine ausreichende visuelle Salienz besitzt.

In einer Reihe von Arbeiten ist informell untersucht worden, welche Merkmalstypen beeinflussen, ob ein Objekt in einer städtischen Umgebung Landmarkencharakter besitzt (z.B. [Lyn60], [ALM64], [CS69]). Weitgehend ungelöst ist die Frage nach qualitativ und quantitativ berechenbaren Kriterien, die a priori angeben, ob ein Objekt eine Landmarke ist. Gleiches gilt für die Verwendung von Referenzobjekten in sprachlichen Beschreibungen (z.B. [Fil71], [Bie67], [MJL76], [Lev89]). Ausgehend von einem merkmalsbasierten Ansatz wird durch eine zweite empirische Untersuchung am Beispiel der beiden Merkmalstypen *Farbe* und *Höhe* überprüft, wie diese die Referenzierung von Objekten beeinflussen. Dabei wird untersucht, in welchen Fällen eine Landmarke referenziert und mit welchen sprachlichen Ausdrücken sie beschrieben wird. Wie Treisman in zahlreichen Experimenten festgestellt hat, ist die Auffälligkeit (Salienz) solcher visueller Merkmale von ihrem Verhältnis zur Umgebung (Kontext) abhängig (vgl. Kapitel 2). Treismans Untersuchungen sind sämtlich in Laboruntersuchungen erhoben worden. Eine Übertragung der Theorie auf reale, makroskopische Phänomenbereiche steht noch aus. Zu diesem Zweck wird in der zweiten Untersuchung analysiert, inwieweit die Merkmalstypen *Farbe*, als Repräsentant eines *rekursiven* Merkmals, und *Höhe* als Repräsentant eines *linearen* Merkmals die Referenzierung einer Landmarke in einer Beschreibung beeinflussen.

Im folgenden werden zwei empirische Untersuchungen vorgestellt. In der ersten empirischen Untersuchung wurden die Probanden aufgefordert, in einer realen Umgebung einem Rezipienten inkrementelle Wegbeschreibungen zu geben. In der zweiten Untersuchung mußten die Probanden Beschreibungen bezüglich einer computeranimierten Kreuzungssituation verbalisieren.

## 3.2 Explorative Untersuchung in realen Umgebungen

Welches raumbezogene Wissen Probanden aufnehmen, wenn sie sich durch städtische Umgebungen bewegen, ist aus unterschiedlichen Perspektiven untersucht worden (vgl. Kapitel 2). Lynch untersuchte als erster, welches raumbezogene Wissen von Einwohnern verschiedener amerikanischer Städte verwendet wird. In ähnlicher Weise wurden in der im folgenden beschriebenen Untersuchung die Probanden aufgefordert, einen Weg schrittweise, während der Navigation durch eine stadtähnliche Umgebung, in möglichst natürlicher Weise zu beschreiben.

### 3.2.1 Arbeitsmodell

Als Arbeitsmodell wird angenommen, daß die Generierung von inkrementellen Wegbeschreibungen eine konventionalisierte, kommunikative Handlung ist. Dies bedeutet, daß sowohl der Sprecher, als auch der Rezipient schematisches Wissen über den prototypischen Ablauf einer inkrementellen Wegbeschreibung besitzen. Dies erlaubt es dem Sprecher, mit minimalem sprachlichen Aufwand die Information zu übermitteln, von der er/sie ausgeht, daß der Rezipient sie benötigt, um den Pfad eindeutig zu finden (vgl. [Kle79], [Wun78], [HG94]). In vollständigen Wegbeschreibungen besteht der größte Aufwand für den Rezipienten darin, die Beschreibungen in ausreichender Weise zu memorieren, um sie später zu verifizieren (vgl. [HCR90]). Da die In-

formation inkrementeller Wegbeschreibungen direkt in der aktuellen Situation verwendet wird, ist zu vermuten, daß inkrementelle Wegbeschreibungen kürzer und prägnanter sind. Weiterhin wird postuliert, daß die Detailliertheit einer Beschreibung von der Zeitdauer bis zum Erreichen eines Verzweigungspunktes abhängt.

### 3.2.2 Durchführung

Die Aufgabe der Probanden bestand darin, entlang eines vorgegebenen Pfads auf dem Gelände der Universität des Saarlandes vom Informatikgebäude bis zur Universitätsbibliothek zu Fuß zu gehen und ihn währenddessen einer begleitenden Person (Rezipient) schrittweise zu beschreiben. Dazu erhielten die Probanden eine Straßenkarte des Universitätsgeländes mit einem eingezeichneten Pfad (siehe Abbildung 3.1). Nachdem der Proband sich auf der Karte und in der Umgebung orientiert hatte, startete er vom Eingang des Informatikgebäudes aus. Der Rezipient ging dem Probanden voraus und hatte keinen Blickkontakt zum ihm, um non-verbale Indikationen des Weges zu vermeiden. Der Rezipient änderte seine Bewegungsrichtung solange nicht, wie der Proband keine entsprechende Anweisung gab. Dieses Verhalten des Rezipienten wurde dem Probanden vorher nicht mitgeteilt. Gleichzeitig nahm der Rezipient den kompletten Pfad und dessen Beschreibung durch eine mitgeführte VHS-Videokamera mit eingebautem Mikrofon auf. Die Zeit für das Abschreiten des Weges vom Informatikgebäude bis zur Universitätsbibliothek betrug zwischen acht bis zwölf Minuten. Es wurden 12 Probanden verschiedenen Geschlechts untersucht, die alle Studenten verschiedener Semesterzahl und unterschiedlicher Fachrichtungen der Universität des Saarlandes waren.

Zur Auswertung sind alle Videoaufzeichnungen transkribiert worden. Dazu wurde der Pfad in 10 Pfadabschnitte aufgeteilt. Der Anfangs- und Endpunkt eines Pfadabschnitts wurde durch einen Verzweigungspunkt festgelegt, an dem der Proband angeben mußte, ob z.B. abgebogen oder eine Treppe hinauf- oder hinabgegangen werden sollte. Bei der Transkription wurde einerseits jede sprachliche Äußerung niedergelegt und andererseits, wieviel Sekunden diese vor dem Erreichen des nächsten Verzweigungspunktes gegeben wurde. Ein Verzweigungspunkt (VP) ist in der untersuchten Domäne durch eine Straßenkreuzung gegeben.

Eine Beispieltranskription, welche zur Beschreibung des Pfadabschnitts 8 gegeben wurde, ist wie folgt:

„... und hier vorne jetzt an der Kreuzung links.\ (10 sec) Abschnitt 8

Dieses Transkript gibt neben der Beschreibung selbst an, daß nach der vollständigen Verbalisierung des Satzes 10 Sekunden vergingen, bis der Rezipient den Verzweigungspunkt am Ende des Pfadabschnitts 9 erreicht hat.

### 3.2.3 Ergebnis

Insgesamt wurden 142 einzelne Beschreibungen aufgenommen. Eine Analyse der Transkripte ergab, daß inkrementelle Wegbeschreibungen aus vier Grundkategorien und einer Restkategorie aufgebaut sind, die unterschiedlich häufig verwendet wurden (vgl. Abbildung 3.2). In die Restkategorie wurden Aussagen, wie z.B. „Vorsicht Stufe!“ aufgenommen. Insgesamt betrug der Anteil der Beschreibungsteile, die in die Restkategorie aufgenommen wurden, 26% der Fälle. Ausdrücke der Restkategorie sind beispielsweise „... so ...“ „... und ...“ sowie „... okay, also ...“. Diese Ausdrücke markieren in den meisten Fällen den Beginn einer neuen Beschreibung bzw. eine Gedankenpause. Da solche Ausdrücke direkt keine räumliche Information enthalten, werden sie im weiteren Verlauf nicht eingehender betrachtet.

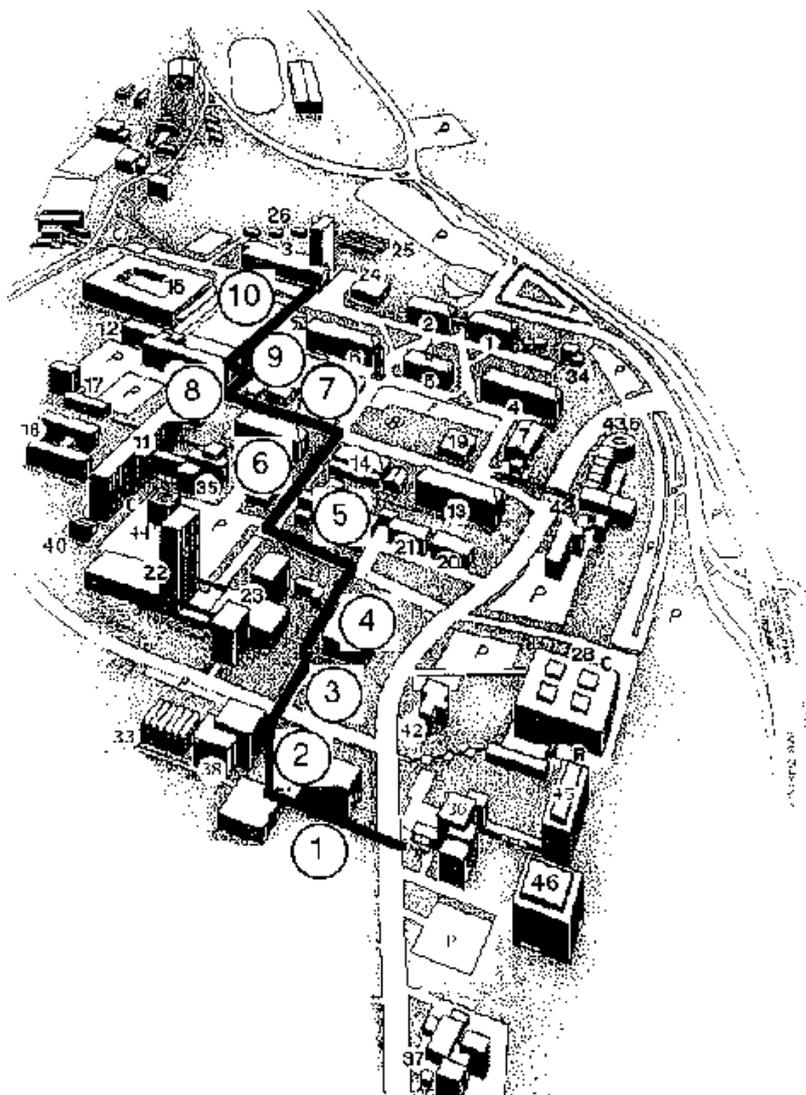


Abbildung 3.1: Kartenausschnitt des Universitätsgeländes mit eingezeichnetem Pfad

In Abbildung 3.2 ist die relative Häufigkeit der relevanten Inhaltskategorien aufgezeichnet.<sup>5</sup> Ein Ausdruck der Kategorie *WOHIN-K* wurde in 39% der Fälle am meisten verwendet, gefolgt von Instantiierungen der Kategorien *WO-K* (33%), *WANN-K* (17%) und *WAS-K* (11%). Die Analyse einzelner Inhaltskategorien ist aber nur bedingt aussagekräftig bezüglich der inhaltlichen Struktur einer inkrementeller Wegbeschreibungen auf Satzebene.

Mehr Aufschluß gibt eine Analyse, wie einzelne Inhaltskategorien miteinander zu komplexeren Beschreibungen kombiniert werden. Es zeigt sich, daß Inhaltskategorien in Abhängigkeit von der verbleibenden Zeit bis zum Erreichen des nächsten Verzweigungspunktes charakteristisch kombiniert werden. Im Korpus sind vorrangig fünf Kombinationen von Inhaltskategorien aufgetreten (vgl. Abbildung 3.3). Mit 37% ist die Kombination *WO-K+WOHIN-K*, gefolgt von der alleinigen Verwendung von *WOHIN-K* am häufigsten gebraucht worden. Zeitliche Referenzen wurden zusätzlich in den anderen drei Kategoriekombinationen integriert. Es liegt die Vermutung nahe, daß sich die relativen Häufigkeiten der kombinierten Inhaltskategorien aus den relativen Häufigkeiten der einzelnen Inhaltskategorien ableiten lassen. Wie eine Analyse jedoch zeigt, ist nur ein

<sup>5</sup>Diese relativen Häufigkeiten sind von den Anteilen der Restkategorie bereinigt.

Abbildung 3.2: Relative Häufigkeiten der relevanten Inhaltskategorien abzüglich der Restkategorie

schwacher Zusammenhang zu erkennen, der darauf beruht, daß kombinierte Inhaltskategorien aus einzelnen Inhaltskategorien zusammengesetzt sind. Beispielsweise läßt sich aus den relativen Häufigkeiten der einzelnen Inhaltskategorien *WO-K* und *WOHIN-K* nicht erklären (vgl. Abbildung 3.2), warum die relative Häufigkeit der Kombination von *WO-K* und *WOHIN-K* größer als die der *WOHIN-K* Kategorie ist (vgl. 3.3).

Als wesentliches Element ist in jeder dieser Kombinationen eine *WOHIN-K* Kategorie enthalten. Daß in diesen Kombinationen selten eine *WAS-K* Kategorie explizit verwendet wird, erscheint dadurch bedingt, daß die Art der Aktion eindeutig determiniert ist. Indikationen einer *WAS-K* Instanz werden häufig nur durch Nennung von Partikeln, wie beispielsweise „ab“ bei „abbiegen“ gegeben.

Eine Analyse der Zeitpunkte, an denen inkrementelle Wegbeschreibungen gegeben worden sind, ergibt eine Häufung von Beschreibungen kurz vor dem Erreichen eines Verzweigungspunktes (Zeitrahmen T<sub>5</sub> in Abbildung 3.4). Je weiter der Proband noch vom Verzweigungspunkt entfernt ist, um so gleichmäßiger sind die Beschreibungen zeitlich verteilt. Je näher der Sprecher

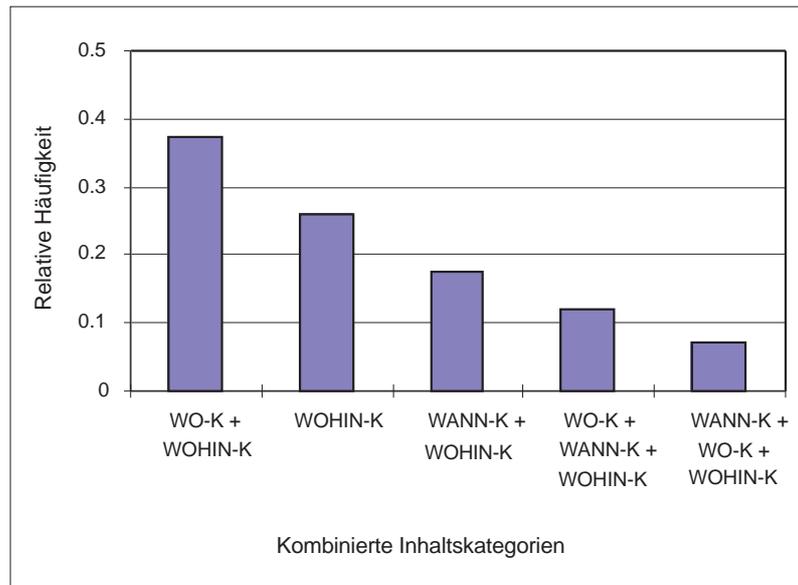
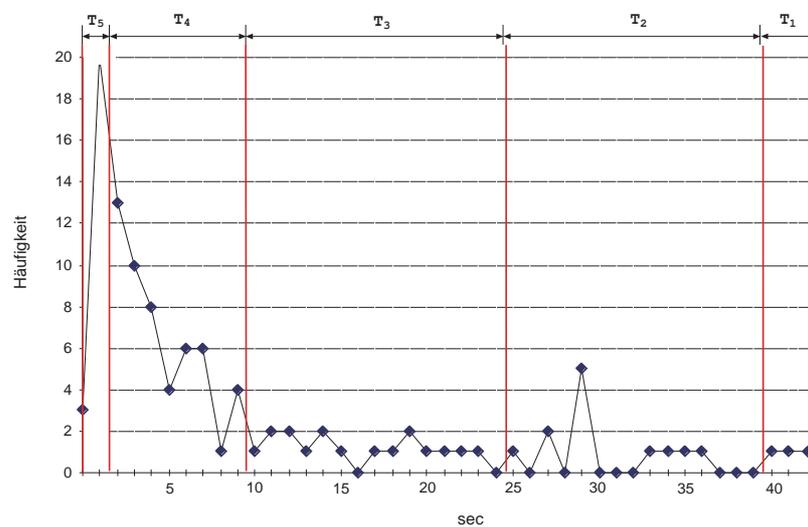


Abbildung 3.3: Relative Häufigkeiten kombinierter Inhaltskategorien

dem nächsten Verzweigungspunkt (VP) ist, um so kürzer werden die Beschreibungen. Inhaltlich wird in einem größeren zeitlichen bzw. räumlichen Abstand zum VP das Gebiet um den VP durch *WO-K* Instanzen genauer beschrieben. Kurz vor Erreichen von VP wird die auszuführende Aktion deiktisch durch eine zeitliche oder räumliche Referenz (vor allem durch „hier“ und „jetzt“) markiert. Der Sprecher geht davon aus, daß das Gebiet um den VP hinreichend genau vom Rezipienten erkannt worden ist.

Abbildung 3.4: Relative Häufigkeiten von Wegbeschreibungen während der Zeitrahmen  $T_i$ 

Sprachlich zeigt sich dieser Wechsel dadurch, daß in einem großen raum-zeitlichen Abstand vollständige, wohlgeformte Sätze gebildet werden, in denen räumliche und zeitliche Referenzen sowie Aktionen genau spezifiziert werden. Bei Annäherung an den Verzweigungspunkt werden nur noch elliptische Beschreibungen gegeben, die zumeist auf deiktischen Referenzen beruhen. Die Analyse der Transkripte bezüglich der Häufigkeiten von inkrementellen Wegbeschreibungen

erbrachte fünf *Zeitraumen*:  $T_1, \dots, T_5$ . Bedingt durch sprachliche und inhaltliche Unterschiede der Beschreibungen zeigt sich, daß die einzelnen Zeitraumen nicht äquidistant sind, sondern um so kürzer werden, je näher sie dem nächsten VP sind. Die nachfolgende Auflistung zeigt Beispiele für Beschreibungen, die den einzelnen Zeitraumen zugeordnet werden.

1.  $T_1$ : „Und vor dem Musikbau da vorne auf der linken Seite gehts rechts die Treppe hoch. Also zwischen Post und Musikgebäude“ (85 sec. bis zum Erreichen des VP)
2.  $T_2$ : „Und am roten Baum da vorne wieder rechts und dann die Treppe hoch“ (35 sec. bis zum VP)
3.  $T_3$ : „... und da an der Kreuzung nach rechts, da so den Berg rauf.“ (15 sec. bis zum VP)
4.  $T_4$ : „... so und hier an der Kreuzung gehts jetzt links.“ (10 sec. bis zum VP)
5.  $T_5$ : „Hier rechts.“ (2 sec. bis zum VP)

Beschreibungen, die sich  $T_1$  zuordnen lassen, rekurren meist auf Erfahrungen der Probanden über die Umgebung. Solche Effekte werden im weiteren nicht weiter behandelt. Beschreibungen innerhalb der Zeitraumen  $T_2$  bis  $T_5$  hingegen sind nahezu ausschließlich durch Referenzen auf visuell perzipierbare Entitäten und Merkmale erfolgt. Nur in fünf Ausdrücken wurde insgesamt auf den Namen eines Gebäudes referenziert. Diese Einteilung in Zeitraumen ist eine erste Approximation. In weiteren Untersuchungen muß analysiert werden, inwieweit es sich um diskrete oder kontinuierliche Prozesse handelt. Eine genaue Analyse zeigt, daß die Probanden, mit Ausnahme von einem, bei dieser Einteilung während der Zeitraumen  $T_3$  bis  $T_5$  eine nahezu identische relative Häufigkeit besitzen, was durch die geringe Standardabweichung bezüglich der einzelnen Zeitraumen von 16% indiziert wird.

Zeitraumen	Dauer bis zum VP (sek)
$T_5$	1-3
$T_4$	3-10
$T_3$	10-25
$T_2$	25-40
$T_1$	> 40

Tabelle 3.1: Zuordnung der Zeitraumen zu Zeitintervallen

Die meisten Beschreibungen werden während des Zeitraumens  $T_4$  gegeben. Diese Tendenz ist, bis auf eine Versuchsperson, typisch. Während  $T_2$  und  $T_3$  differieren die Häufigkeiten zwischen den Personen stark, was daran liegt, daß die Beschreibung des Gebiets um den VP in unterschiedlicher Weise erfolgte. Beschreibungen zum Zeitraumen  $T_2$  sind sehr ausführlich, wobei während  $T_3$  bereits eine Fokussierung erfolgt.

Eine Zuordnung des Inhalts einer Beschreibung zu den Zeitraumen ist durch Rückgriff auf die Inhaltskategorien möglich. Dabei zeigt sich, daß die Häufigkeit von Instanzen einzelner Inhaltskategorien während den Zeitraumen  $T_2$  und  $T_3$  in der gleichen Größenordnung liegen (vgl. Abbildung 3.5). Während des Zeitraumens  $T_4$  kommt es zu einem parallelen Anstieg in der Verwendung von Instanzen der *WANN-K* und *WAS-K* Kategorien. Im Zeitraumen  $T_5$  geht die Verwendung von Instanzen der beiden Kategorien wieder auf ein Minimum zurück. In ähnlicher Weise verläuft die Verwendung von Instanzen der *WOHIN-K* und der *WO-K* Kategorien parallel. Nach einem Minimum im Zeitraumen  $T_3$  steigt die Häufigkeit an und erreicht in  $T_5$  ein Maximum. Dies indiziert, daß Instanzen der *WANN-K* und *WAS-K* Kategorien dazu verwendet werden, Aktionen

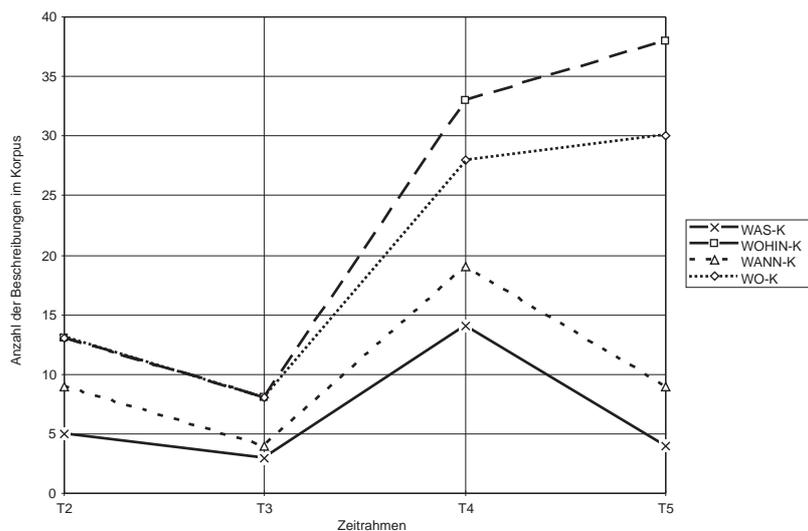


Abbildung 3.5: Zeitbezogene Häufigkeiten von Inhaltskategorien

frühzeitig anzukündigen. In dem Augenblick, in dem die Aktion ausgeführt werden soll, wird die räumliche Lokation (*WO-K*) zusammen mit einer Richtungsangabe (*WOHIN-K*) beschrieben.

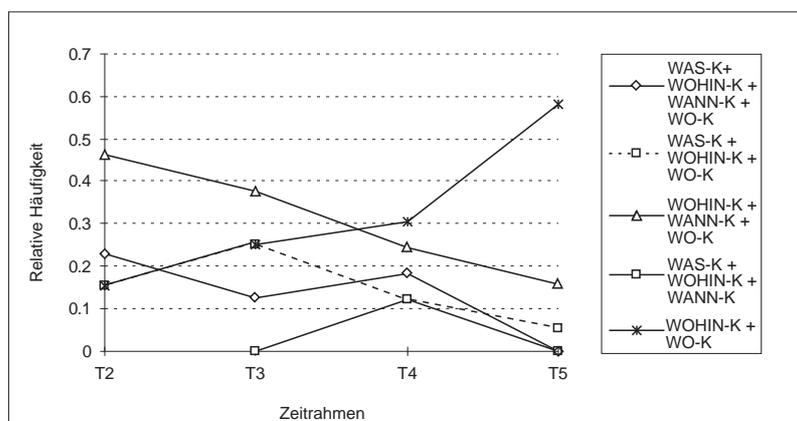


Abbildung 3.6: Relative Häufigkeiten der Kombination von Inhaltskategorien bezüglich einzelner Zeitraumen

Die Verwendungshäufigkeit einzelner Kategorien hat Auswirkungen auf die Inhaltsstruktur der gesamten Beschreibung. In Abbildung 3.6 sind die relativen Häufigkeiten von Kategorienkombinationen dargestellt. Prinzipiell lassen sich zwei Tendenzen identifizieren: Je kürzer die Zeit bis zum Erreichen eines Verzweigungspunktes ist, um so stärker steigt die relative Häufigkeit, mit der die *WOHIN-K + WO-K* Kombinationen verwendet werden. Ein inverses Verhalten zeigt sich für die Verwendung der anderen Kombinationen, die auf Erweiterungen von *WAS-K + WOHIN-K* Kombinationen beruhen und ebenso für die Kombination *WOHIN-K + WANN-K + WO-K*.

	WAS-K	WOHIN-K	WANN-K	WO-K	relative Häufigkeit
$T_2$	•	•	•	•	0,23
	•	•		•	0,15
		•	•	•	0,46
		•		•	0,15
$T_3$		•		•	0,25
	•	•		•	0,25
		•	•	•	0,38
	•	•	•	•	0,13
$T_4$	•	•	•	•	0,18
	•	•		•	0,12
		•	•	•	0,24
		•		•	0,30
	•	•	•		0,12
		•	•		0,03
$T_5$		•		•	0,58
	•	•			0,05
		•			0,07
	•	•		•	0,05
		•	•	•	0,15
		•	•		0,07

Tabelle 3.2: Zeitrahmenbezogene relative Häufigkeiten von  $W^*$ -Kategoriekombinationen

### 3.2.4 Diskussion

Inkrementelle Wegbeschreibungen bestehen, bis auf eine Restkategorie, aus räumlichen, zeitlichen und aktionsorientierten Referenzen. Aus diesen Elementen lassen sich Inhaltskategorien ableiten, mit denen die transkribierten Beschreibungen der Untersuchung analysiert worden sind. Wie sich herausstellte, werden einzelne Inhaltskategorien in charakteristischer Weise miteinander kombiniert. Bedingt durch die Dynamik, die durch die Eigenbewegung des Sprechers induziert wird, werden Beschreibungen zeitabhängig gegeben. Somit läßt sich einerseits die Schemahaftigkeit vollständiger Wegbeschreibungen (vgl. Kapitel 2.5.2) auch bei inkrementellen Wegbeschreibungen nachweisen. Andererseits sind inkrementelle Wegbeschreibungen stark durch zeitliche Restriktionen geprägt. Instanzen von Inhaltskategorien werden dabei in flexibler Weise in Abhängigkeit von der verbleibenden Zeit bis zum Erreichen des nächsten Verzweigungspunkts kombiniert. Jede Kombination enthält eine *WOHIN-K* Kategorie, wohingegen die anderen Kategorien optional verwendet werden. Dies läßt sich dadurch erklären, daß die *WOHIN-K* Kategorie der Träger der wesentlichen Information bezüglich Wegbeschreibungen ist, die durch andere Kategorien optional unterstützt wird. Die Aufgabe der anderen Kategorien ist es, Aktionen (*WAS-K*), räumliche Lokationen (*WO-K*) oder temporale Referenzen (*WANN-K*) zu beschreiben, damit der Rezipient diese Elemente eindeutig identifizieren kann. Besonders in zeitkritischen Situationen wird die Angabe der Richtung durch einen raumdeiktischen Ausdruck in einer räumlichen Lokation verankert (*WOHIN-K+WO-K*). Die primäre Aufgabe einer inkrementellen Wegbeschreibung ist es, die Richtungsänderung einer Aktion an einem Verzweigungspunkt zu indizieren. Je näher der Sprecher und der Rezipient dem nächsten VP sind, um so stärker werden richtungsindizierende Instanzen der *WOHIN-K* Kategorie verwendet. Eine Identifikation einer Aktion erfolgt in zeitkritischen Situationen meist durch Angabe räumlich-qualitativer Beschreibungen von Richtungsänderungen, die in räumlichen Lokationen verankert werden. Je

mehr Alternativen zu unterscheiden sind, von denen der Beschreiber annimmt, daß sie den Hörer irritieren können, um so genauer wird diese durch Verkettung von *WO-K* Instanzen räumlich verankert (vgl. [Man86]). Eine elaborierte Beschreibung des räumlichen Gebietes um den Verzweigungspunkt erfolgt nur, wenn dem Beschreiber ausreichend viel Zeit bleibt (während  $T_2$  und  $T_3$ ).

Die Aufteilung in Zeitrahmen korrespondiert mit der Unterscheidung in verschiedene subjektive Räume (z.B. [Itt73], [Man83], [GG87], [Mon93]). Obschon der Proband die aktuelle Situation in  $T_2$  und  $T_3$  überblickt, befindet er sich noch innerhalb des rein *visuellen Raums*. Entsprechend der Einteilung in Zeitrahmen erstreckt sich der visuelle Raum von  $T_2$  über  $T_3$  bis hin zu  $T_4$ . Mit Übergang zum Zeitrahmen  $T_5$  erfolgt ein Übergang vom *visuellen Raum* zum *interaktiven Raum* (Handlungsraum). Dieser Übergang manifestiert sich in der häufigen Verwendung deiktischer Referenzen. Vollständige Wegbeschreibungen auf der Basis räumlicher Langzeitrepräsentationen des Sprechers basieren auf einem ähnlichen Übergang (vgl. [Kle79], [Wun78]). Da der Übergang nicht auf perzipierter Information basiert, sondern imaginiert erfolgt, korrespondiert dies mit dem Konzept der *imaginären Wanderung* (vgl. [LL75]). Der Sprecher deutet dabei dem Rezipienten durch Verwendung von deiktischen Beschreibungen an, daß das Gebiet erreicht worden ist, an dem eine Aktion erfolgt. Dies bewirkt, daß der Rezipient seine Aufmerksamkeit auf das beschriebene Gebiet richtet, um die verbalsprachlich gegebenen Referenzen mit geeigneten Referenten zu identifizieren und die Aktion auszuführen. Dies erfolgt typischerweise durch redundante Beschreibungen eines einzelnen Sachverhalts. Eine solche Verwendung redundanter Beschreibungen zeigt sich auch in dem hier aufgebauten Korpus der inkrementellen Wegbeschreibungen. Dabei zeigt eine Analyse, daß redundante Beschreibungen charakteristisch verteilt gegeben werden. Wird eine Beschreibung zum Zeitrahmen  $T_i$  gegeben, so wird eine redundante Beschreibung zum Zeitrahmen  $T_{i+2}$  gegeben.

Es lassen sich fünf nicht äquidistante Zeitrahmen unterscheiden. Dies wird einerseits durch die sprachliche und andererseits durch die inhaltliche Struktur der Beschreibungen belegt. Inwieweit diese fünf Zeitrahmen eine erste Approximation einer generellen Einteilung des subjektiven Raums darstellen oder aber feiner eingeteilt werden müssen, sollte Gegenstand weiterer Untersuchungen sein. Den Kategoriekombinationen lassen sich relative Häufigkeiten in Abhängigkeit von den Zeitrahmen zuordnen (vgl. Tabelle 3.2). Je nachdem, welche Kategorieentypen in einer Situation instantiiert werden können, wird eine Beschreibung auf der Basis unterschiedlicher Kategoriekombinationen generiert.

Weiterhin ist zu untersuchen, welchen Einfluß die Variation der Bewegungsgeschwindigkeit des Probanden auf die Verwendung von Inhaltskategorien und deren Kombinationen besitzt. Insbesondere im Hinblick auf die Analyse von Ressourcenbeschränkungen können solche Untersuchungen Hinweise liefern.

### 3.3 Untersuchung in einer computeranimierten 3D-Umgebung

In Beschreibungen visuell wahrgenommener, räumlicher Umgebungen referenziert ein Sprecher nicht auf alle Details. Dies entspricht vor allem kognitiven und kommunikativen Ökonomiekriterien (vgl. [Gri75]). Eine Teilaufgabe, die ein Sprecher während der inhaltlichen Festlegung einer inkrementellen Wegbeschreibung zu lösen hat, besteht darin, solche Elemente aus einer Umgebung auszuwählen, die eindeutig identifizierbar sind. Eine solche referenzierende Rolle übernehmen *Landmarken* (vgl. Kapitel 2.3.4). In makroskopischen, stadtähnlichen Umgebungen werden häufig Gebäude als Landmarken verwendet (z.B. [Lyn60], [ALM64]). Daß Landmarken eine wichtige Rolle beim Erwerb und der Beschreibung raumbezogenen Wissens spielen, ist vielfach diskutiert worden (vgl. Kapitel 2.3.4). Offen bleibt in diesen Arbeiten, auf Grund welcher Kriterien ein Objekt als Landmarke verwendet wird. Gleiches gilt für situationssemantische Mo-

delle, in denen zwischen sprachlichen Ausdrücken und referenzierten Objekten Verbindungen (“connection”) postuliert werden (vgl. [BP83]).

In der vorhergehend beschriebenen Untersuchung ist analysiert worden, wie inkrementelle Wegbeschreibungen sprachlich strukturiert und adaptiv verwandt werden. Da Landmarken in realen Umgebungen nicht systematisch verändert werden können, ist in einer weiteren Studie untersucht worden, wie visuelle Merkmale die Verwendung eines Objektes als Landmarke beeinflussen.<sup>6</sup> Basierend auf Anne Treismans “feature integration theory” werden exemplarisch das *rekursive* Merkmal *Farbe* und das *lineare* Merkmal *Höhe* untersucht (vgl. Kapitel 2). Beide Merkmale sind bereits im Hinblick auf visuelle Suchmechanismen untersucht worden (z.B. [Lap66]). Bei der Beschreibung einer Landmarke werden solche Merkmale bevorzugt, die, entsprechend einer geeigneten Metrik, eine maximale Distanz zu den entsprechenden Merkmalen der Kontextobjekte aufweisen (vgl. [Man86]). Empirisch ist belegt, daß Farbe einfacher als Größe und diese einfacher als Form zu bestimmen ist (vgl. [Man86]). Bei der Verwendung von Merkmalen in Beschreibungen semantisch armer Umgebungen determinieren sprachstrukturelle Eigenschaften von Merkmalspezifikationen die Wahrscheinlichkeit einer Nennung von Objektmerkmalen (vgl. [Pob91]). Aber nicht jedes Merkmal, welches leicht zu erkennen ist, ist gleichzeitig auch leicht zu beschreiben.<sup>7</sup> Der Frage nach der Verwendung der Merkmale *Farbe* und *Höhe* in inkrementellen Wegbeschreibungen bei der Referenzierung von Landmarken wird vor diesem Hintergrund in der folgenden Untersuchung nachgegangen.

### 3.3.1 Arbeitsmodell

Das Arbeitsmodell, welches der empirischen Untersuchung zugrunde liegt, postuliert, daß sowohl durch visuelle Merkmale, als auch durch räumliche Lokationen beeinflusst wird, ob ein Objekt in einer inkrementellen Wegbeschreibung als Landmarke verwendet wird. Damit verbunden ist die Frage, wodurch die unterstützende Referenzierung visueller Merkmale beeinflusst wird. Bezogen auf die einzelnen visuellen Merkmale wird vermutet, daß das Merkmal *Farbe* den größten Einfluß besitzt (vgl. [Man86]). Weiterhin wird postuliert, daß die visuelle Salienz eines Objektes sich aus einer gewichteten Linearkombination visueller Merkmale zusammensetzt, und daß die Verwendung eines Objektes als Landmarke von pfadbasierten Intentionen abhängt.

### 3.3.2 Durchführung

Die Untersuchung wurde computergestützt in einem Labor durchgeführt, bei dem der Proband direkt vor einem Bildschirm saß. Es nahmen dieselben Personen teil, die zuvor an der ersten Untersuchung teilgenommen hatten. Da diese Untersuchung vor der in Unterkapitel 3.2 beschriebenen explorativen Untersuchung gemacht worden ist, sind Interferenzen zwischen beiden Untersuchungen weitgehend auszuschließen. Die Probanden bekamen die Aufgabe in einer Animation einer Kreuzungssituation einen Abbiegevorgang zu beschreiben, der ihnen zuvor in Form eines Pfeiles auf dem Bildschirm angezeigt wurde. Die Kreuzungssituation bestand aus zwei Straßen, die sich in einem Winkel von 90° schnitten und vier Objekten, die auf den Ecken der Kreuzung positioniert waren (vgl. Abbildung 3.7).

Die Untersuchung bestand aus vier Teilen (vgl. Tabelle 3.3). In den ersten beiden Teilen wurde

---

<sup>6</sup>In dieser Untersuchung ist der Fall nicht betrachtet worden, daß Landmarken beschrieben werden, um die Lokation einer Kreuzung von anderen Kreuzungen hervorzuheben. In diesem Falle stellt sich die Frage, wie ein primärdeiktischer Raum unabhängig von der intendierten Bewegungsrichtung etabliert wird. Da die Richtung der Bewegung am nächsten Verzweigungspunkt nur eine untergeordnete Rolle spielt, ist zu vermuten, daß zur Identifikation der Kreuzung auch Landmarken auf der „falschen“ Seite verwendet werden.

<sup>7</sup>Eine eingehende Diskussion der allgemeinen Wirkung von Merkmalen auf Klassifikations- und Sprachprozesse findet sich in [Kli71].

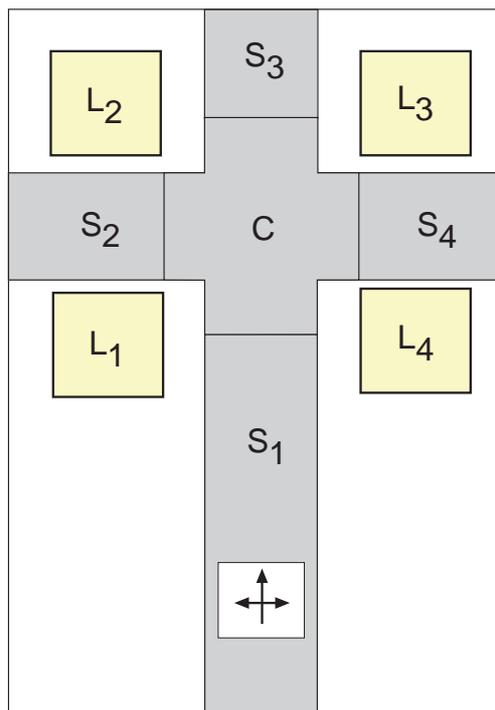


Abbildung 3.7: Kreuzungssituation mit den Pfadabschnitten  $S_i$ , der Kreuzung  $C$  und den Gebäuden  $L_i$

	Freier Fall	Festgelegter Fall
Farbe	Gruppe 1	Gruppe 2
Höhe	Gruppe 2	Gruppe 1

Tabelle 3.3: Zuordnung von Probandengruppen zu Untersuchungsvariationen

die Farbe genau einer Landmarke variiert und in den anderen beiden deren Höhe. Um überprüfen zu können, inwieweit sprachliche Phänomene bei der Referenzierung von Landmarken wesentlich sind, wurde die Variation der *Farbe* und der *Höhe* in zwei Teile geteilt. Im ersten konnten die Probanden die Situation frei beschreiben (freier Beschreibungsfall), wohingegen sie im zweiten Teil aufgefordert wurden, eine bestimmte, über eine Nummer indizierte Landmarke in die Beschreibung zu integrieren (festgelegter Beschreibungsfall).

Die Probanden wurden in zwei Gruppen à sechs Personen geteilt. Die erste Gruppe beschrieb die Situation mit variierenden Farben im freien Modus und mit variierender Höhe im festgelegten Modus. Im Gegensatz dazu beschrieb die zweite Gruppe Situationen mit variierenden Höhen im freien und variierende Farben im festgelegten Modus.

Für beide Variationen erhielten die Probanden die Instruktion, sich vorzustellen, als Beifahrer in einem Auto zu sitzen.<sup>8</sup> Ihre Aufgabe bestand darin, so natürlich wie möglich einem neben ihnen sitzenden Rezipienten verbal zu beschreiben, wohin dieser zu fahren hat. Der Proband wurde darüber informiert, daß kurz vor Beginn der animierten Situation am unteren Rand ein Pfeil eingeblendet wird, der die Richtung der Aktion am nächsten Verzweigungspunkt angibt. Ein Pfeil nach oben indizierte eine Geradeausfahrt, ein Pfeil nach links einen Linksabbiegevor-

<sup>8</sup>Da die Probanden sich nicht bewegen sollten, ist diese Anweisung natürlicher, als wenn sie sich vorstellen, ein Fußgänger zu sein.

Farbe	H/L/S 1	H/L/S 2	H/L/S 3	H/L/S 4
grau	000/075/000	000/065/000	000/055/000	000/045/000
rot	029/052/058	029/042/065	029/032/085	029/025/100
blau	180/054/014	180/044/015	180/034/019	180/024/017

Tabelle 3.4: Tabelle der variierten Farbwerte (H), Helligkeiten (L) und Sättigungsgrade (S)

gang, respektive ein Pfeil nach rechts einen Rechtsabbiegevorgang. Durch Tastendruck konnte der Proband die jeweils nächste Situation starten. Im festgelegten Fall mußten die Probanden zusätzlich die Landmarke in die Beschreibung integrieren, deren Nummer am unteren Rand des Bildschirms erschien. Die Referenz der Nummer auf die Landmarke wurde durch eine Karte hergestellt, die den Probanden vorlag (vgl. Abbildung 3.7).

Die graphischen Animationen der Kreuzungssituation sind in LISP mit CLIM-Funktionen implementiert worden (vgl. Kapitel 4.3.2). Das Animationsprogramm auf der Basis der Programmpakete *GEOMETRIX2* und *GEODISPLAY* (vgl. Kapitel 4.3) lief auf einer Workstation von Hewlett Packard (HP 9000/720) und zeigte die Animationen auf einem 19"-Farbmonitor (Hewlett Packard HP98754A) an. Der Proband interagierte mit dem Animationssystem über den Druck der Leerzeichen-Taste.

### 3.3.2.1 Methode: Farbvariation

Im ersten Fall der Variation des Farbmerkmals eines Objektes wurden die Farbmerkmale der anderen Objekte konstant hell-grau (HLS<sup>9</sup>: 000/065/000) gehalten. Alle Objekte hatten identische geometrische Ausmaße. Variiert wurde der Farbwert, die Helligkeit und der Sättigungsgrad, die Lokation und die Richtung der zu beschreibenden Aktion.<sup>10</sup> Als Farbgrundwerte wurden exemplarisch *grau*, *rot* und *blau* mit vier verschiedenen Helligkeits- und Sättigungswerten verwendet (vgl. Tabelle 3.4).

### 3.3.2.2 Ergebnis: Farbvariation

Wurde ein Objekt im freien Beschreibungsfall sprachlich referenziert, erfolgte die Nennung des Farbmerkmals über Farbwerte, Helligkeits- und Sättigungsvariationen, Lokationen des Objektes und Aktionen hinweg im Durchschnitt mit einer relativen Häufigkeit von 14% (Standardabweichung von 5% über alle Farbvariationen). Im festgelegten Beschreibungsfall kam es zu einer hohen Standardabweichung von 17% über alle Farbvariationen. Im Durchschnitt lag die Verwendung der Farbmerkmals zur Beschreibung einer Landmarke bei 32% (vgl. Abbildung 3.8).

Durchschnittlich wurde im freien Beschreibungsfall mit gleicher Häufigkeit eine Landmarke ohne (10% der Fälle) und mit Indikation ihres Farbmerkmals (14% der Fälle) referenziert. Im festgelegten Falle hingegen wurde in 19% der Fälle häufiger eine Landmarke ohne (51%) als mit Nennung des Farbmerkmals (32%) beschrieben. Diese Unterschiede blieben über Farbwerte und Helligkeits- und Sättigungsgradvariationen hinweg konstant (vgl. Abbildung 3.8).

Ein deutlicher Unterschied in der Referenzierung des Farbmerkmals ergibt sich beim Vergleich der Beschreibungen von Abbiegeaktionen und Geradeausfahrten. Bei Geradeausfahrten wurde im freien Beschreibungsfall keine Landmarke durch ihr Farbmerkmal genauer spezifiziert. Im festgelegten Fall hingegen wurde in Beschreibungen mit Objektreferenzen bei einer Geradeausfahrt

<sup>9</sup>hue, lightness, saturation

<sup>10</sup>Ein spezieller Farbabgleich ist nicht durchgeführt worden, da es in erster Linie um die Untersuchung qualitativer, nicht quantitativer, Unterscheidungen ging.

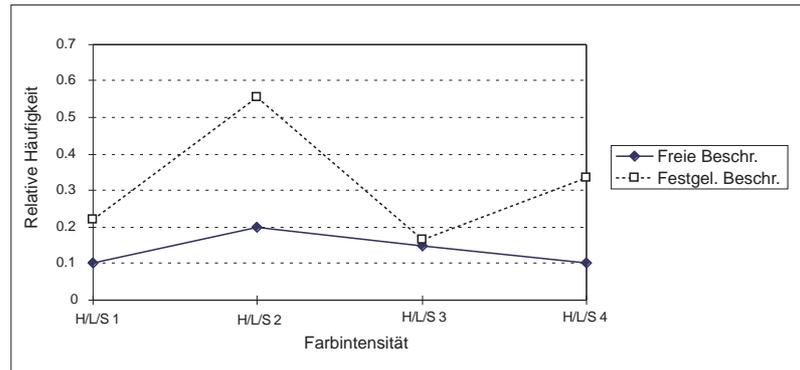


Abbildung 3.8: Relative Häufigkeitsverteilung für das Merkmal *Farbe* im freien und im festgelegten Beschreibungsfall bei unterschiedlicher Farbintensität

mit einer Häufigkeit von 46% auf das Farbmerkmal referenziert. Dies ändert sich grundsätzlich bei intendierten Abbiegeaktionen. Im freien Fall wurden in 73% der Beschreibungen mit Landmarken (Standardabweichung 9%), bzw. in 25% aller Beschreibungen (Standardabweichung 3%) das Farbmerkmal genannt. Im festgelegten Fall verringerte sich dies auf 48% (Standardabweichung 6%). Dies zeigt, daß Probanden vor allem bei Abbiegeaktionen Farbmerkmale mit ausreichender visueller Salienz referenzierten. Der festgelegte Fall zeigt, daß zwar ein Farbmerkmal bei einer Geradeausfahrt auffällig ist, es jedoch nie verwendet wird.

Die räumliche Lokation der Landmarke ist in den Beschreibungen von herausragender Bedeutung. Waren die Probanden frei in ihrer Beschreibung, so verwandten sie für die Beschreibung einer Abbiegeaktion ausschließlich (100%) solche Landmarken, deren Lokation auf der Abbiege-seite („richtige“ Seite, vgl. Abbildung 3.9) lag. Auch wenn eine Landmarke auf der zur Aktion abgewandten Seite („falsche“ Seite) farblich besonders auffällig war, wurde sie in der Beschreibung nicht berücksichtigt.

**3.3.2.2.1 Diskussion: Farbvariationen** Bezogen auf farblich auffällige Objekte zeigt sich, daß das Zusammenspiel aus visueller Salienz, räumlicher Lokation eines Objektes und den Intentionen des Sprechers determiniert, ob eine Landmarke in eine Beschreibung integriert wird. Bei einem Abbiegevorgang mit Referenz einer Landmarke wurde in 73% aller Beschreibungen die Landmarke durch ihr Farbmerkmal genauer spezifiziert. Dieser Effekt ist dabei weitgehend vom Farbwert und den assoziierten Helligkeits- und Sättigungsgradvariationen unabhängig. Im festgelegten Fall sinkt diese Verwendungshäufigkeit auf 45% ab. Eine mögliche Interpretation ist, daß die Probanden das Farbmerkmal als derart salient betrachteten, daß sie davon ausgingen, das diese Landmarke alleine durch den visuellen Eindruck eine geringe Ambiguität besitzt. Eine solche Schlußfolgerung ist mit der Analyse anderer empirischer Arbeiten konform (vgl. [Man86]). Auf Grund kognitiver Ökonomie wird in solchen Fällen das Farbmerkmal nicht gesondert referenziert.

Von großem Einfluß auf die Verwendbarkeit einer Landmarke in einer Aktionsbeschreibung ist deren räumliche Lokation. Der Sprecher integriert bei einer Abbiegeaktion ausschließlich Landmarken, die auf der „richtigen“ Seite liegen. Da Landmarken in dieser Versuchsreihe immer in konstantem Abstand zum Verzweigungspunkt positioniert waren, ist weiter zu untersuchen, inwieweit andere Parameter, wie Distanzen und Winkelabweichungen, die Referenzierbarkeit einer Landmarke beeinflussen.

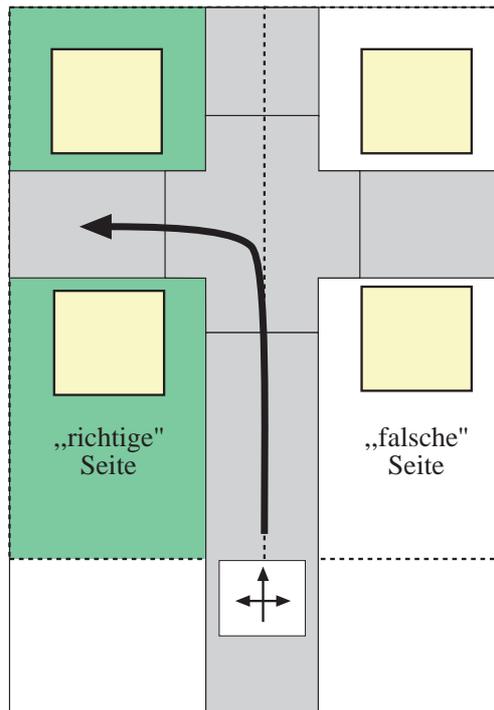


Abbildung 3.9: Qualitative Unterscheidung in Landmarken, die bei einem Linksabbiegevorgang auf der „richtigen“ (A und B) und solche, die auf der „falschen“ Seite (X und Y) liegen

### 3.3.2.3 Methode: Höhenvariation

Bei Variation der Höhenmerkmale eines Objektes wurde die Höhe der anderen drei Objekte konstant gehalten (15 Einheiten). Alle Objekte hatten identische neutrale Farbmerkmale (H/L/S: 000/065/000). Variiert wurde der Höhenwert, die Lokation und die Richtung der Aktion. Die Höhenwerte wurden im Verhältnis 1 : 1,3 : 1,6 : 2 zur Standardhöhe variiert. Die gesamte Situation lag zu Anfang einer Animationssequenz vollständig im Blickfeld des Probanden.

### 3.3.2.4 Ergebnis: Höhenvariation

Wurde eine Landmarke im freien Beschreibungsfall sprachlich referenziert, erfolgte die Nennung des Merkmals *Höhe* über Höhenwert-, räumliche Lokations- und Aktionsvariationen hinweg im Durchschnitt mit 17% bei einer Standardabweichung von 17% über alle Höhenvariationen. Diese hohe Streuung ist dadurch bedingt, daß die Referenzierung des Höhenmerkmals mit ansteigendem Höhenwert bis auf 36% erheblich ansteigt. Im festgelegten Fall wurde ein Höhenmerkmal in 51% der Fälle bei einer Standardabweichung von 4% über alle Höhenvariationen verwendet (vgl. Abbildung 3.10). Dies zeigt, daß das Merkmal *Höhe* im festgelegten Fall über alle Höhenvariationen hinweg gleich verwendet wird. Im freien Beschreibungsfall wurde eine Landmarke in nur 2% aller Fälle referenziert und dabei in 17% dieser Fälle zusätzlich durch ihr Höhenmerkmal genauer spezifiziert. Im festgelegten Fall verschob sich dieses Verhältnis auf 5% zu 51%.

Ein signifikanter Unterschied zur Referenzierung des Farbmerkmals ergibt sich beim Vergleich der Beschreibungen von Abbiegeaktionen und Geradeausfahrten. Bei Geradeausfahrten wurde im freien Beschreibungsfall keine Landmarke durch ihr Höhenmerkmal genauer spezifiziert. Im festgelegten Fall wurde relativ zu allen Beschreibungen über Höhenwerte hinweg bei einer Geradeausfahrt konstant mit einer Häufigkeit von 89% auf das Höhenmerkmal referenziert.

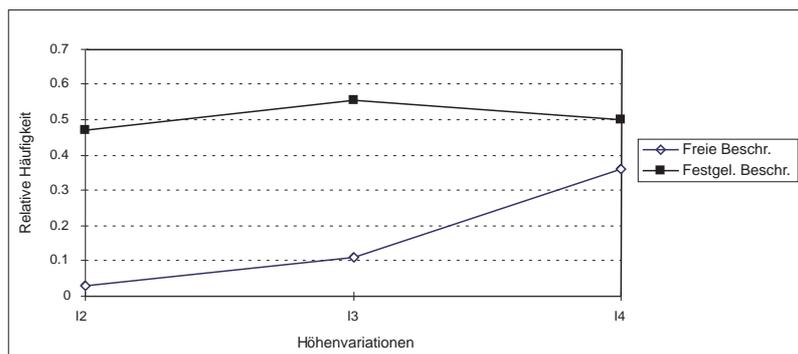


Abbildung 3.10: Relative Häufigkeiten für das Merkmal *Höhe* im freien und im festgelegten Beschreibungsfall bei verschiedenen Höhen  $I_i$

Dies ändert sich deutlich bei intendierten Abbiegeaktionen. Im freien Fall wurden in 53% aller Beschreibungen (Standardabweichung 46%) bzw. in 93% der Beschreibungen, in denen Landmarken (Standardabweichung 11%) referenziert wurden, das Höhenmerkmal genannt. Die große Streuung kommt wiederum dadurch zustande, daß über die Höhenwerte hinweg die Referenzierung des Höhenmerkmals von 8% bis auf 100% zunimmt. Im festgelegten Fall erhöht sich die Referenzierungshäufigkeit auf 97% (Standardabweichung von 5%).

Der identische Effekt wie bei der Variation von Farbmerkmalen gilt auch hier für die räumliche Lokation der variierten Landmarke. Waren die Probanden frei in ihrer Beschreibung, so verwendeten sie für die Beschreibung eines Abbiegevorgangs ausschließlich (100%) Landmarken, die auf der „richtigen“ Seite lagen. Auch wenn eine Landmarke auf der „falschen“ Seite besonders hoch war, wurde sie in einer Beschreibung nicht berücksichtigt.

**3.3.2.4.1 Diskussion: Höhenvariation** Aus den Ergebnissen folgt, daß bei einem Abbiegevorgang mit Referenz einer Landmarke in 93% aller Beschreibungen sie durch ihr Höhenmerkmal genauer spezifiziert wurde. Dieser Effekt ist dabei weitgehend davon unabhängig, wie hoch die Landmarke ist. Vielmehr mußte die Höhe nur einen Schwellwert überschreiten. Im festgelegten Fall steigt die Verwendungshäufigkeit auf 97% an. Ebenso wie in der vorhergehenden Untersuchung des Farbmerkmals ist für die Verwendbarkeit einer Landmarke in einer Aktionsbeschreibung deren räumliche Lokation von großem Einfluß. Der Sprecher integriert ausschließlich Landmarken, die auf der „richtigen“ Seite bezüglich einem Abbiegevorgang liegen.

Bei einer Geradeausfahrt spielt im freien Beschreibungsfall das Höhenmerkmal keine Rolle. Wurde der Proband jedoch aufgefordert, diese Landmarke zu integrieren, so wurde sie in 89% aller Fälle und in 97% bei einem Abbiegevorgang zusätzlich durch ihr Höhenmerkmal spezifiziert.

### 3.3.3 Gesamtdiskussion

Die Ergebnisse der empirischen Untersuchungen zu Farb- und Höhenmerkmalen von Landmarken zeigen, daß das Zusammenspiel aus visueller Auffälligkeit eines Merkmals, räumlicher Lokation des Objektes und den Intentionen des Sprechers determinieren, ob eine Landmarke in die Beschreibung integriert wird. Die Variation des Merkmals ergibt nur für das Höhenmerkmal einen klaren Effekt. Bei Farbmerkmalen muß die Auffälligkeit einen Schwellenwert überschreiten, so daß ein qualitativ sichtbarer Unterschied feststellbar wird, damit sie für eine genauere Spezifikation einer Landmarke verwendbar sind.

Einen starken Effekt besitzen pfadbasierte Intentionen des Abbiegens (links- und rechtsabbie-

gen), wohingegen bei einer Geradeausfahrt eine Landmarke selten referenziert wird. Es läßt sich schließen, daß Geradeausfahrten keiner weiteren raumbezogenen Spezifikation bedürfen und als Standardfall angenommen werden. Dies ist dadurch begründet, daß die Aktion des Geradeausfahrens räumlich gesehen eindimensional verläuft und nur eine geringe Winkelvariation zwischen dem aktuellen und dem nächsten Pfadabschnitt zu berücksichtigen ist. Im Gegensatz dazu werden Abbiegeaktionen bezüglich eines zweidimensionalen Raums definiert, wodurch eine genauere Beschreibung notwendig ist, um den intendierten Pfad eindeutig zu referenzieren.

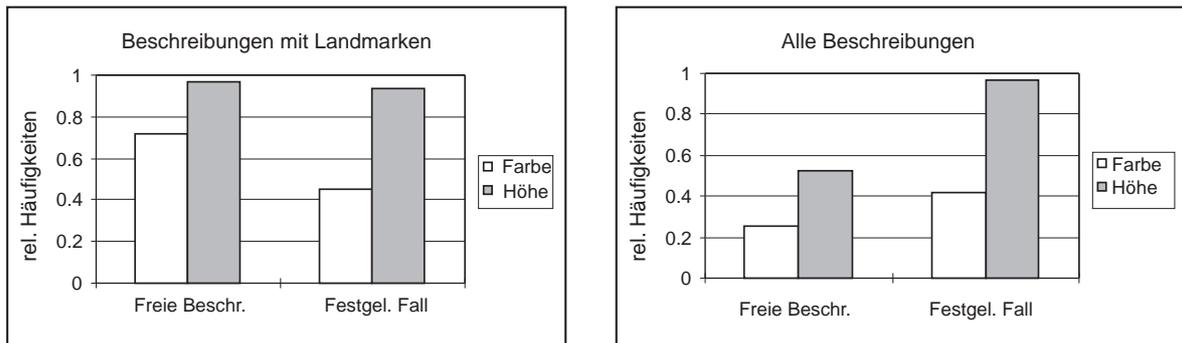


Abbildung 3.11: Vergleich der relativen Häufigkeiten der Referenzierungen von Höhen- und Farbmerkmalen a) relativ zu Beschreibungen mit Landmarken und b) relativ zu allen Beschreibungen

Bei einem Abbiegevorgang werden Landmarken fast grundsätzlich verwendet. Besitzt eine referenzierte Landmarke ein auffälliges Farbmerkmal, so wird sie in 73% aller Fälle durch dieses genauer spezifiziert. Bei einem auffälligen Höhenmerkmal erfolgt dies in 93% der Fälle. Der Referenzierung einer Landmarke kommt eine gebietskonstituierende Bedeutung zu. Damit der Rezipient dieses Gebiet eindeutig identifizieren kann, versucht der Sprecher, die Landmarke so genau wie möglich zu beschreiben. Es werden aber grundsätzlich nur solche Landmarken referenziert, die auf der „richtigen“ Seite liegen. In den Fällen, in denen die Probanden Landmarken auf der „falschen“ Seite integrieren sollten, hatten sie erhebliche Probleme. Bei einem Probanden führte dies sogar dazu, daß er nicht in der Lage war, die Aktion rechtzeitig zu beschreiben. Dies ist ein weiteres Indiz für die Schemahaftigkeit der Generierung von Wegbeschreibungen (vgl. [Kle79], [Wun82], [RS87]).

Das Verhältnis zwischen der Verwendung von visuellen Merkmalen *Höhe* und *Farbe* bei einem Abbiegevorgang zeigt im Vergleich, daß das Merkmal *Höhe* auf sprachlicher Ebene von stärkerem Einfluß ist als *Farbe*, was im Widerspruch zur Erkennbarkeit dieser Merkmale steht (vgl. [Man86]). Im freien und im festgelegten Beschreibungsfall ergibt sich, daß ein Höhenmerkmal doppelt so häufig sprachlich spezifiziert wird, wie ein Farbmerkmal (vgl. Abbildung 3.11). Dies unterstützt eine Asymmetrievermutung zwischen visueller Wahrnehmung und sprachstruktureller Umsetzung visueller Merkmale (vgl. [Man86]).



## Kapitel 4

# Formale Beschreibung des kognitiven Wegbeschreibungsagenten

Bei der Generierung von inkrementellen Wegbeschreibungen wird auf eine Vielzahl kognitiver Fähigkeiten rekurriert. Das Szenario der inkrementellen Beschreibung von Wegen während der Navigation durch eine unbekannte Umgebung bedingt u.a. den Einsatz sprachlicher, perzeptueller und navigationsbezogener Fähigkeiten. Es wird in diesem Kapitel ein agentenorientiertes, formales Modell eines *kognitiven Agenten* vorgestellt, dessen Aufgabe es ist, inkrementelle Wegbeschreibungen bezüglich einer simulierten dreidimensionalen Umgebung zu generieren. Dazu wird zwischen einer *funktionalen* und einer *prozeßorientierten Beschreibungsebene* unterschieden. Das funktionale Modell determiniert den generischen Rahmen, der durch verschiedene Prozeßmodelle instantiiert werden kann. Eine mögliche Instanz ist das nachfolgend beschriebene, implementierte Prozeßmodell MOSES (siehe Kapitel 4.3).

Der Agent ist ein *dynamischer Entscheidungsträger*, der Eingabeparameter mit internen Systemzuständen integriert, um ein adaptives Verhalten zu generieren. Unter adaptivem Verhalten wird verstanden, daß der Agent sein Verhalten an Eingabeparameter anpaßt. Die Eingabeparameter ändern sich dynamisch aufgrund der Situation, und durch das Verhalten des Agenten. Von den Eingabeparametern und dem internen Zustand des Agenten werden Ressourcenbeschränkungen abgeleitet, welche determinieren, was für ein Verhalten ein Agent potentiell zeigen kann. Unter den potentiellen Verhaltensweisen wählt der Agent eine maximal *adäquate* aus. Eine solche Betrachtungsweise integriert die Sicht *reaktiver* und *planbasierter* Systeme. Stark reaktiv geprägt sind die Funktionen der visuellen Objektauswahl. Jedoch sind sie von den Intentionen und den Zielen des Agenten nicht unabhängig. Die Bestimmung von Inhaltsstrukturen und die Generierung von sprachlichen Beschreibungen fallen unter intentionale und dadurch planbasierte Verhaltensweisen. Solche Funktionen können aber nicht losgelöst von Änderungen in den Situationen betrachtet werden, so daß sie in enger Beziehung zu reaktiven Funktionen stehen. Der hier vorgestellte Ansatz eines kognitiven Wegbeschreibungsagenten ist weder rein planbasiert, noch rein reaktiv zu verstehen. Vielmehr wird eine *verhaltensorientierte Sichtweise* eingenommen, die danach fragt, wie einzelne Funktionen integriert werden können, um ein bestimmtes Verhalten zu erzeugen.

Entsprechend dieser Fragestellung wird bezüglich des Phänomens der inkrementellen Wegbeschreibungen ein funktionales Agentenmodell vorgestellt, welches (synthetische) visuelle Daten, Karteninformation, navigationsrelevante Information und Annahmen über den Hörer in zeitliches und räumliches Wissen umsetzt, um das sprachliche Verhalten des Agenten daran anzupassen. Für den Agenten stellen sich in erster Linie *Auswahlprobleme*. Die Information, die ihm durch eine Situation und eine Karte zur Verfügung steht, ist weitaus umfangreicher als das, was er zum Erfüllen seiner Aufgabe benötigt. Das Auswahlproblem wirft die Frage auf, wie ein

Agent „die richtige Information“ aus der Situation und internen Wissensstrukturen auswählt. Dies korreliert mit Kleins Übergang von einem *Primär-* zu einem *Sekundärplan* (vgl. Kapitel 2.5.2). Im Gegensatz dazu erfolgt der Zugriff auf raumbezogenes Wissen nicht über eine Kognitive Karte sondern wird von visuell perzipierter Information abgeleitet. Hierfür müssen Kriterien definiert werden, mittels derer der Agent entscheidet, was „das Richtige“ in einer gegebenen Situation ist. Dies gilt auf der Perzeptionsseite und auf der Präsentationsseite. Es soll nur „das Richtige“ aufgenommen, aber auch nur „das Richtige“ präsentiert werden. Nicht immer ist ohne weiteres formalisierbar, was „das Richtige“ ist (vgl. [RW91], [DB88]). Besonders, wenn Allgemeinwissen eine Rolle spielt, läßt sich das Problem nur näherungsweise formalisieren. Aus diesem Grunde wird in diesem Modell, ebenso in vergleichbaren Modellen, die Berücksichtigung von Allgemeinwissen auf ein Minimum reduziert.

Die Beschreibung des funktionalen Modells des Agenten untergliedert sich in verschiedene Betrachtungsebenen und Teilphänomene. Der Agent integriert Funktionen zur *Evaluation von Kartenmaterial* und *visueller Information*. Auf der Ausgabeseite verfügt er über Funktionen zur *natürlichsprachlichen Generierung* von inkrementellen Wegbeschreibungen und zur *Steuerung der Navigation*. Jedem dieser Teilphänomene ist auf der funktionalen Ebene ein *Modul* zugeordnet. Über ein Modul werden Funktionen zu einer Einheit integriert, die, wenn sie in einer Laufzeitumgebung instantiiert sind, als *Prozesses* bezeichnet werden.

Um ein komplexes Verhalten, wie die Generierung und Präsentation von inkrementellen Wegbeschreibungen zu zeigen, müssen die beteiligten Module in koordinierter Weise miteinander interagieren. Eine solche Interaktion wird vom sogenannten *Steuerungsmodul* übernommen. Seine Aufgabe ist es, entsprechend den situationsabhängigen Ressourcenbeschränkungen, durch sogenannte *Phasen*, andere Module zu aktivieren bzw. zu deaktivieren. Phasen explizieren das koordinierte Zusammenspiel zwischen Modulen und integrieren diese zu komplexen Verarbeitungsverbänden. In einem Modul werden wiederum *Funktionen* in koordinierter Weise dazu verwendet, eine dem Modul gestellte Aufgabe zu lösen. Funktionen sind die grundlegenden Bestandteile für die Beschreibung des formalen Modells. Aus der Gesamtheit der Funktionen und Repräsentationen lassen sich alle potentiellen Verhaltensweisen ableiten. Welches Verhalten real gezeigt wird, hängt von der Instantiierung des funktionalen Modells und der Laufzeitumgebung ab.

Der Aufbau dieses Kapitels ist wie folgt. Zuerst wird der modulare Aufbau des Modells beschrieben. Anschließend werden die Module im einzelnen vorgestellt. Zum Abschluß dieses Kapitels werden die Funktionen der jeweiligen Module eingehend diskutiert.

## 4.1 Modularer Aufbau des funktionalen Modells

Dem funktionalen Modell des Agenten liegt eine Teilung in sechs Module zugrunde (vgl. Abbildung 4.1). Auf der Eingabe- und Ausgabeseite handelt es sich um vier *periphere* Module: das *visuelle Objektauswahlmodul* (VOA-M), das *Pfadsuchmodul* (PS-M), das *Sprachproduktions- und präsentationsmodul* (NLP-M) sowie das *Navigationsmodul* (NV-M). Zwischen diesen Modulen liegt das *Raumintegrationsmodul* (RI-M), in dem raumbezogenes Wissen aus verschiedenen Quellen integriert und gespeichert wird. Gesteuert werden diese Module durch das *Steuerungsmodul* (K-M), dessen Aufgabe die Kontrolle sowie die Adaption des Agenten an *Ressourcenbeschränkungen* ist.

Das visuelle Modul wählt, basierend auf physikalischen, visuellen Merkmalen aus einer Situation visuell saliente Objekte aus. Das Modul zur Pfadsuche bestimmt durch heuristische, inkrementelle Strategien die zu verfolgenden Pfadabschnitte. Auf der Ausgabeseite übernimmt das Sprachproduktionsmodul die Aufgabe, relevantes Wissen zu versprachlichen und dieses zu ei-

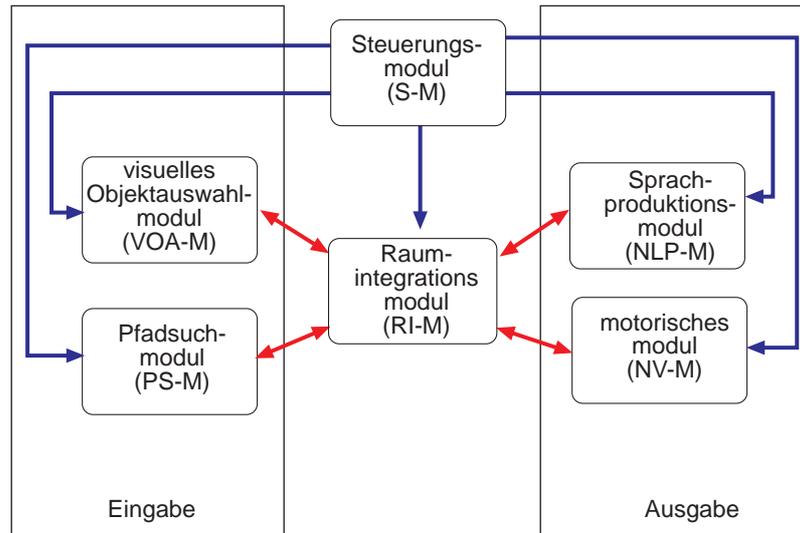


Abbildung 4.1: Module des funktionalen Modells

nem *geeigneten* Zeitpunkt zu präsentieren. Das Navigationsmodul umfaßt Funktionen, mit denen Bewegungen und damit allgemein die Navigationsfähigkeiten des Agenten auf einer abstrakten Ebene verbunden sind. Es ergibt sich eine vernetzte, funktionale Struktur (siehe Abbildung 4.1). Das Steuerungsmodul koordiniert die Aktivierungen und Deaktivierungen von Modulen und nimmt dadurch eine zentrale Stellung ein. Im folgenden wird zuerst das Steuerungsmodul (Kapitel 4.1.1), dann die Eingabemodule (Kapitel 4.1.2,4.1.3), das Raumintegrationsmodul (Kapitel 4.1.4) und abschließend das Ausgabemodul (Kapitel 4.1.5,4.1.6) beschrieben.

#### 4.1.1 Das Steuerungsmodul

Das Steuerungsmodul hat die Aufgabe, den *Kontrollfluß* zwischen den Modulen zu kontrollieren. Dies erfolgt vor allem auf der Grundlage der zur Verfügung stehenden zeitlichen Ressourcen. Zeitliche Ressourcen werden aus Evaluation der räumlichen Struktur, der aktuellen Situation und der raumzeitlichen Beziehungen des Agenten bestimmt. Dazu schätzt das Steuerungsmodul, wieviel Zeit zur Bestimmung und Ausführung einer Handlung zur Verfügung steht. Ein Überschreiten von zeitlichen Restriktionen muß, bezogen auf inkrementelle Wegbeschreibungen, grundsätzlich vermieden werden, da andernfalls der Agent Gefahr läuft, eine Beschreibung zu spät oder gar nicht zu präsentieren. Zeitliche Restriktionen im Kontext inkrementeller Wegbeschreibungen sind demzufolge *harte* Ressourcenbeschränkungen, die notwendigerweise eingehalten werden müssen, um die Bedingung für *adäquates* Verhalten zu erfüllen. Die Abschätzung zeitlicher Ressourcenbeschränkungen wird durch eine Funktion zur Evaluation zeitlicher Beziehungen (*Evaluate-Resource-Time*) übernommen. Die Aufgabe dieser Funktion besteht darin, die durch die Eigenbewegung des Agenten und durch die Situation vorgegebenen raumzeitlichen Beschränkungen miteinander in Beziehung zu setzen, so daß für den Agenten eine Abschätzung für den zur Verfügung stehenden Zeitrahmen ( $T_{\text{guess}}$ ) bestimmt werden kann. *Evaluate-Resource-Time* führt dadurch die Transformation der *externen* Ressource Zeit in eine *interne* durch.<sup>1</sup>

Das Steuerungsmodul nimmt eine prominente Stellung im Modell ein. Die einzige Steuerung, die das Steuerungsmodul ausübt, ist die *Aktivierung* und *Deaktivierung* von Modulen entsprechend

<sup>1</sup>In Fällen, in denen der Agent diese Internalisierung nicht vollführt, kann man umgangssprachlich davon sprechen, daß der Agent „die Zeit vergessen“ hat.

einem *Ablaufplan*. Ein Ablaufplan definiert, in welcher Reihenfolge Phasen aktiviert werden. Sie können als kompliziertes Wissen verstanden werden, welches dem Agenten erlaubt, sich in komplexen Situationen effizient zu verhalten. Maßgeblich für diese Kontrolleinflüsse ist das zeitliche Verhalten der einzelnen Module. Im Falle, daß ein Modul eine Aufgabe nicht im vorgegebenen Zeitintervall erfüllt, wird dieses Modul deaktiviert. Zur Strukturierung der Aktivierung von Modulen dienen sogenannte *Phasen*. Ein Phase legt fest, in welcher Reihenfolge Module bezüglich einem Teilproblem aktiviert werden. Phasen werden zwecks Erreichen kohärenten Verhaltens aller Module zu *Ablaufplänen* integriert. In Kapitel 4.2.1 werden die Funktionen des Steuerungsmoduls vorgestellt.

#### 4.1.2 Das Pfadsuchmodul

Verfügt ein Akteur über kein ausreichendes räumliches Langzeitwissen bezüglich einer Umgebung, ist er auf externe Informationsquellen angewiesen. Liegt das Zielobjekt nicht innerhalb des Perzeptionsbereichs, ist es eine Standardkulturtechnik, Straßen- bzw. Umgebungskarten zu verwenden (vgl. [DL88]).

Eine häufig verwendete Vorgehensweise zur Modellierung des Kartenlesens ist es, räumliches Wissen aus einer abstrakten, räumlichen Repräsentationsstruktur zu extrahieren, wobei der Zusammenhang zwischen Repräsentation und dem zu repräsentierenden weitgehend unbeachtet bleibt. Nur in wenigen Modellen wird explizit der Zusammenhang zwischen der räumlichen Information, die durch die Umgebung oder Artefakte bereitgestellt wird und der räumlichen Repräsentationsstruktur eines Agenten hergestellt (vgl. Kapitel 2.6).

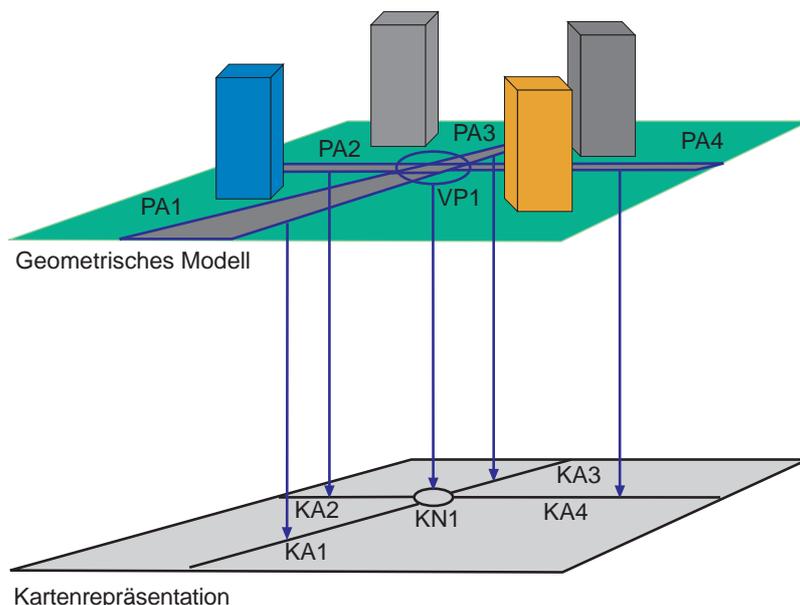


Abbildung 4.2: Transformation eines geometrischen 3D-Modells in eine 2D-Straßenkartenrepräsentation

Das integrierte Pfadsuchmodul wählt heuristisch und inkrementell aus einer Kartenrepräsentation bezogen auf eine Start- und Ziellokation Pfadabschnitte aus. Es wird ein zweistufiger Ansatz verwendet. Auf der Ebene der geometrischen Repräsentation einer Umgebung sind Pfadabschnitte und Verzweigungspunkte definiert. Diese werden auf einer Kartenrepräsentationsebene, die in Form von nichtgerichteten Graphen gegeben ist, abgebildet. Entsprechend den Ergebnissen über Segmentierungseffekte bei Pfaden (siehe Kapitel 2.3.5), besteht eine direkte Abbildung zwischen

*Pfadabschnitten* (vgl. Definition 4) und *Kanten* eines Graphen (vgl. Abbildung 4.2). Die Definition eines Pfadabschnitts ist subjektiv auf den Betrachter bzw. Agenten bezogen. Dies trägt dem Umstand Rechnung, daß verschiedene Betrachter einen Pfad verschieden einteilen können. *Knoten* eines Graphen repräsentieren *Verzweigungspunkte*, an denen der Agent seine Orientierung ändert (vgl. Definition 5). In gleicher Weise wie ein Pfadabschnitt ist ein Verzweigungspunkt relativ zum Betrachter bzw. Agenten definiert. Dadurch ist die Festlegung eines Verzweigungspunktes ebenso wie die eines Pfadabschnitts eine empirische Frage. Neben der Änderung des egozentrischen Referenzsystems spielen auch funktionale, soziale und physikalische Faktoren bei der Konstituierung eines Verzweigungspunktes in einer Umgebung eine Rolle. Dies eröffnet ein weites Feld für die empirische Untersuchung von Segmentierungseffekte von Pfaden. Im weiteren Verlauf dieser Arbeit wird sich auf Verzweigungspunkte im Sinne der Definition 5 beschränkt.

Die Lagebeziehung zwischen einem Knoten, der die aktuelle Lokation und dem Knoten, der die Ziellokation repräsentiert, konstituiert Konfigurationswissen, welches dem Agenten erlaubt, die grobe Richtung einzuhalten. Solches Wissen wird, bedingt durch die Vogelperspektive beim Lesen einer Karte, bezüglich eines globalen, allozentrischen Referenzsystems etabliert. Auf lokaler Ebene ist der Agent dadurch in der Lage abzuschätzen, ob ein Pfadabschnitt oder eine Lokation zu weit von einer idealen Verbindung abweicht. Die Lagebeziehung von aneinandergrenzenden Pfadabschnitten, die in einer Situation identifiziert werden können, sind bezüglich eines egozentrischen Referenzsystems etabliert. Entsprechend Kleins Analyse vollständiger Wegbeschreibungen, wird auch bei inkrementellen Wegbeschreibungen zwischen einer vollständigen und einer inkrementellen Pfadsuchstrategie unterschieden. Im ersten Fall determiniert der Sprecher den Pfad mittels *imaginärer Wanderung* vollständig, bevor er den Weg beschreibt. Im zweiten Fall erfolgen Pfadsuche und Beschreibung abwechselnd (vgl. Kapitel 2.5.2). Diese beiden Strategien sind Extrempole.<sup>2</sup> In Kapitel 4.2.2 werden Funktionen des Pfadsuchmoduls diskutiert, die diese Eigenschaften modellieren.

---



---

#### Definition 4: Ein Pfadabschnitt

Ein *Pfadabschnitt* ist ein vom Betrachter bzw. Agenten als Einheit empfundenes Segment eines Pfades. Auf einem Pfadabschnitt bleibt die Ausrichtung bzw. die Änderung der Ausrichtung des egozentrischen Referenzsystems konstant. Die Größenordnung der Länge eines Pfadabschnittes ist domänenabhängig.

---



---



---



---

#### Definition 5: Ein Verzweigungspunkt

Ein *Verzweigungspunkt* ist eine Lokation, an dem ein Betrachter bzw. Agent die Orientierung bzw. die Änderung der Orientierung des egozentrischen Referenzsystems verändert.

---



---

Grundlegend für die Definition von Pfadabschnitten und Verzweigungspunkten ist die Orientierung bzw. die Änderung der Orientierung des egozentrischen Referenzsystems des Agenten. Hierbei spielt die Domäne eine wesentliche Rolle. Die Etablierung eines Verzweigungspunktes ist

---

<sup>2</sup>Informelle Untersuchungen indizieren, daß zwischen diesen beiden Strategien gewechselt werden kann, bzw. daß die Länge der Strecke, die vorausgeplant wird, variabel ist.

durch die Festlegung eines Pfadabschnitts determiniert. Ein Verzweigungspunkt wird im einfachsten Falle durch das Ende eines linearen Pfadabschnitts festgelegt (siehe Abbildung 4.3). Jedoch können auch Pfadabschnitte, die Kurven enthalten, als Einheit identifiziert werden.<sup>3</sup> Auf einem sich schlängelnden Pfad wechselt die Ausrichtung des egozentrischen Referenzsystems permanent. Ist der Agent in der Lage, die globale räumliche Struktur und das Muster der Änderung der Orientierung zu erkennen, so werden diese Kurven bzw. die zu verfolgende Trajektorie zu einem Pfadabschnitt zusammengefaßt. Ist er dazu nicht in der Lage, so bleibt nichts anderes übrig, als kleine Pfadabschnitte zu definieren, die von einer Änderung der Orientierung zur nächsten reichen.

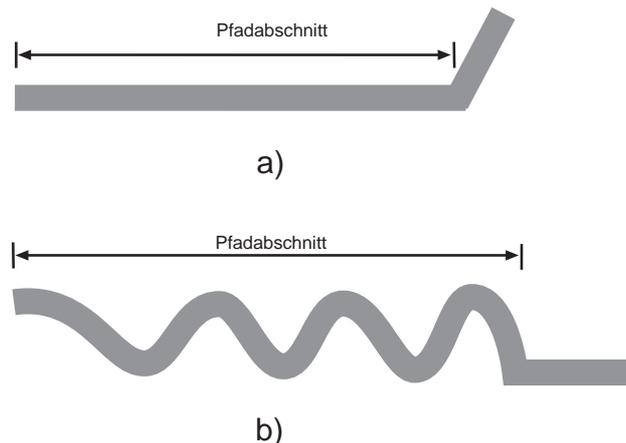


Abbildung 4.3: a) linearer Pfadabschnitt, b) nicht-linearer Pfadabschnitt

Das Ende eines Pfadabschnitts ist durch einen *Verzweigungspunkt* markiert. In manchen Fällen drückt dieser nur einen für diesen Pfadabschnitt untypischen Orientierungswechsel aus. Im allgemeinen aber verläßt der Agent an einer solchen Lokation die Eindimensionalität eines Pfadabschnitts. Alternative Pfadabschnitte werden in verschiedene Richtungen verfolgt. Unter diesen wird in heuristischer Weise der Pfad ausgewählt, welcher bezüglich einer geeigneten Qualitätsmetrik der kostengünstigste ist.

### 4.1.3 Das Objektauswahlmodul

Eine inkrementelle Wegbeschreibung setzt eine Vorstellung über die räumliche Repräsentation einer Situation voraus. Eine solche räumliche Repräsentation integriert Objekte durch Etablierung von räumlichen Relationen (vgl. Kapitel 4.2.4.3). Wird ein Objekt beschrieben, so wird zwischen dem perzipierten Objekt und der Beschreibung eine Referenzbeziehung aufgebaut (vgl. Kapitel 2.5). Objektbeschreibungen werden häufig in visuell perzipierten Objekten verankert. Jedoch eignen sich nicht alle Objekte einer Situation als Referenzobjekte, sondern nur solche, die die Qualität der Beschreibung, wie z.B. durch Verbesserung der Disambiguierungsfähigkeit, erhöhen. Bei visuell verankerten Referenzbeziehungen basiert die Auswahl geeigneter Objekte vor allem auf der Analyse visueller Merkmale. Dies bedingt Funktionen, welche von den sichtbaren Objekten einer Situation solche auswählen, die visuell auffällig sind. Über das Konzept der *visuellen Salienz* läßt sich ein Maß für die Vergleichbarkeit von visuell sichtbaren Objekten definieren. Welche Objekte in einer Situation salient sind, ist zusätzlich vom intentionalem Wissen des Agenten beeinflusst, so daß sowohl reaktive als auch planbasierte Verarbeitungsstrategien zu berücksichtigen sind.

<sup>3</sup>Ein dahingehend zu überprüfendes Kriterium wäre das der *Stetigkeit*.

In bisherigen Ansätzen zur sprachlichen Beschreibung visueller Daten, geht man davon aus, daß immer eindeutig vorgegeben ist, welche Objekte in die Beschreibung eingehen sollen (z.B. [NN83b], [ABHR86]). In Arbeiten zur Bestimmung von räumlichen und sprachlich-räumlichen Relationen werden Referenzobjekte und zu lokalisierende Objekte meist explizit angegeben, so daß das Problem der Objektauswahl weitgehend umgangen wird (z.B. [OMT94], [Gap94]). Ein Ansatz zur Auswahl von Referenzobjekten findet sich in [Gap96].

Das Teilproblem der Auswahl von *Referenzobjekten* ist für einen sich bewegenden Agenten irreführend. „Auswahl“ impliziert, daß eine Situation vollständig analysiert worden ist, bevor die Salienz von Objekten determiniert wird. Interessiert man sich dafür, wie visuelle Daten in dynamischen Kontexten sprachlich beschrieben werden, kann man aufgrund von Ressourcenbeschränkungen nicht von einer vollständigen Analyse der Situation ausgehen. Die visuelle Analyse einer Situation ist eng an die Anzahl der in einer Situation vorkommenden Objekte gebunden (vgl. [Tso85], [Tso86]). Befunde zeigen, daß der Mensch Aufmerksamkeitsmechanismen dazu verwendet, um sich ohne vollständige Analyse der gesamten Situation auf die Antizipation einer Untermenge der angebotenen Information zu beschränken (vgl. Kapitel 2).

Allgemein ist zu untersuchen, welche Funktionen es erlauben, nur die „richtige“ visuelle Information zu analysieren. Diese Frage ist ad hoc gesehen zirkulär, da ein Betrachter in einer Situation nur dann ein Objekt als salient bezeichnen kann, wenn vorher ein Mindestmaß an Analyse in die gesamte visuell-zugängliche Information investiert worden ist. Eine andere Strategie wäre ein wahlloses Durchforsten der visuellen Information. Es wird in dieser Arbeit ein alternativer Ansatz vorgestellt, in dem eine Situation auf einer Merkmalsebene minimal analysiert wird, um Bereiche zu finden, in denen visuell saliente Referenzobjekte zu finden sind.

Mit der Frage nach der Auswahl von visueller Information assoziiert sind formale Modelle, in denen die visuelle Wahrnehmung als *bottom-up* oder *top-down* Verfahren modelliert wird. Erstere sind datengetriebene Modelle. Top-down-Verfahren hingegen sind wissensbasiert in dem Sinne, daß sie den Kontext, die Ziele des Agenten und sein langfristig gespeichertes Wissen berücksichtigen. Das Problem eines reinen *bottom-up* Verfahrens liegt in der Orientierungslosigkeit, da es bei der Analyse relevant ist, wozu der Agent die Information benötigt. Andererseits bewirkt eine *top-down* Analyse, daß in manchen Situationen nach Entitäten und Gesamtzusammenhängen gesucht wird, die in der gegebenen Situation nicht vorhanden sind.

Generell ist visuelle Salienz kein monolithischer Begriff, sondern drückt die visuelle Auffälligkeit eines Objektes in einer bestimmten Situation unter benutzer- und situationsabhängigen Voraussetzungen aus. Sie wird einerseits durch die visuelle Information der Situation und andererseits durch die Ziele und das Wissen des Agenten bestimmt. In dem hier vorgeschlagenen Modell wird angenommen, daß das perzeptive System eine Situation bereits in Objekte zerlegt hat. Für jedes dieser Objekte entscheidet der Agent, welche visuelle Salienz dieses für ihn aus der Perspektive in der aktuellen Situation besitzt. Formal bedeutet dies, daß eine partielle Ordnung über die ausgewählten Objekte einer Situation bestimmt wird. Diese partielle Ordnung kann nicht unabhängig vom Kontext betrachtet werden.<sup>4</sup> Das visuelle System wählt Objekte u.a. nach physikalischen Merkmalen, wie *Farbe*, *Form* und *Bewegungsrichtung* aus. Dies täuscht aber nicht darüber hinweg, daß auch die *Bedeutung* von Objekten für die Bestimmung des Salienzwertes zu berücksichtigen ist. Die Schwierigkeit, die sich dabei ergibt, ist, daß sich die Bedeutung eines Objektes mit der Situation und den Intentionen des Betrachters ändern kann.<sup>5</sup> Für die Bestimmung der Salienz eines Objektes sind demzufolge zumindest *physikalische* und *semantische* Merkmale zu unterscheiden. Physikalische Merkmale sind nur von ihrer Verteilung

<sup>4</sup>Z.B. fällt ein roter Punkt unter vielen blauen auf, wohingegen er unter vielen anderen roten unauffällig ist und somit an Salienz für den Betrachter verliert.

<sup>5</sup>Beispielsweise hat ein Glas voll mit Wasser für einen durstigen Menschen in der Wüste eine andere Bedeutung als in einem Weinkel.

in der Situation abhängig. In alltäglichen Situationen unterliegen physikalische Merkmale auch einer Veränderung ihrer Saliens, wenn ihnen nicht-physikalische Konnotationen zugeordnet werden.<sup>6</sup> In gleicher Weise können Intentionen des Betrachters den Salienswert eines physikalischen Merkmals verändern, wenn er beispielsweise a priori beabsichtigt, nach einem Objekt mit einem bestimmten physikalischen Merkmal zu suchen. Aus diesem Grund erscheinen alle Theorien zur Bestimmung der Saliens eines Objektes oder dessen Merkmale als zu kurzgreifend, wenn der Kontext der Situation und die Intentionen des Betrachters nicht betrachtet werden. Visuelle Saliens eines Objektes kann nur im Zusammenhang mit den Intentionen des Betrachters und der Situation bestimmt werden.<sup>7</sup> Wie sich in den Untersuchungen zu inkrementellen Wegbeschreibungen herausstellte (vgl. Kapitel 3), bewirkt eine Voreinstellung der Probanden auf eine städtische Umgebung, daß vor allem Gebäude und große Kunstobjekte als wesentliche Landmarken ausgewählt wurden.<sup>8</sup> Zu vermuten ist, daß auch andere Objektklassen, wie beispielsweise Schilder und Bäume Landmarkenfunktionen “übernehmen können. Da sie aber im vorliegenden Korpus nicht referenziert worden sind, werden sie o.B.d.A. nicht weiter betrachtet.

Die Integration der Funktionen des Objektauswahlmoduls verwendet eine Mischform aus “bottom-up” und “top-down” Strategie. Ausgehend von globalem Wissen werden physikalische Merkmale von Landmarken im aktuellen Blickfeld dazu verwendet visuelle Saliens zu bestimmen. Eine Situation wird, der “feature integration theory” folgend, über eine Projektion in visuelle Merkmale zerlegt. Die Merkmale der Objekte werden entsprechend ihrem Merkmalsstyp auf topographierhaltende, zweidimensionale Repräsentationen sagittal zur Perzeptionsrichtung des Agenten parallel projiziert. Auf diesen Repräsentationen werden in Abhängigkeit vom Merkmalstyp visuelle Saliens von Objekten bestimmt. Diese einzelnen Projektionen werden übereinandergelegt, um entsprechend der gesamten Klasse der betrachteten Typen die Saliens der Objekte zu berechnen. Der Bereich der *räumlichen Aufmerksamkeit* wird durch *pfadgeleitete Intentionen* des Betrachters gesteuert. Das Ergebnis des Auswahlprozesses ist eine partielle Ordnung über visuell sichtbare Landmarken. Pfadbezogene Intentionen des Agenten veranlassen ihn, nicht auf alle Objekte im Sichtbereich seine räumliche Aufmerksamkeit gleichmäßig zu verteilen, sondern auf bestimmte Bereiche zu fokussieren. Der Fokus ist dabei durch einen Ausschnitt des Sichtfeldes bestimmt, der als *räumlicher Fokus* bezeichnet wird. Der räumliche Fokus ist ein Filter für visuelle Daten im Hinblick auf deren Verwendung in nachfolgenden Funktionen.<sup>9</sup>

In Kapitel 4.2.3 werden Funktionen diskutiert, die merkmalsbasiert visuelle Saliens von sichtbaren Objekten aus einer perspektivischen Sicht heraus bestimmen.

---

<sup>6</sup>Im Straßenverkehr hat die Farbe *rot* eine funktionale Bedeutung und wird dadurch viel eher fokussiert als in anderen Situationen.

<sup>7</sup>Ansätze, die dies nicht berücksichtigen, führen dazu, daß eine gegebene räumliche Anordnung von Objekten in unterschiedlichen Situationen und bei unterschiedlichen Intentionen immer die gleiche partielle Ordnung mit konstanten Salienswerten liefert (siehe u.a. den Ansatz von [GKS89], [Gap96]).

<sup>8</sup>In nicht-städtischer Umgebung übernehmen andere physikalische Entitäten die Funktion einer Landmarke, wie beispielsweise Flüsse, Bäume oder Felsformationen (vgl. [Sch92]).

<sup>9</sup>Rückschlüsse auf pfadgeleitete Intentionen, mittels derer der räumliche Aufmerksamkeitsbereich gesteuert wird, geben die Beschreibungen der empirischen Untersuchungen (siehe Kapitel 3). Besonders interessante Einblicke erlaubten solche Beschreibungen, die unter der Vorgabe gegeben wurden, bestimmte Landmarken an einer Kreuzung zu integrieren. War die Landmarke auf der Abbiegeseite, so gab die Testperson ohne zu zögern typische Wegbeschreibungen. Lag sie jedoch auf der gegenüberliegenden Seite, hatten die Testpersonen Schwierigkeiten, überhaupt eine Beschreibung zu generieren. Es wäre eine zu grobe Vereinfachung anzunehmen, daß dies auf visuelle Funktionen zurückzuführen ist. Dennoch kann festgestellt werden, daß es für den Beschreiber ungewöhnlich ist, Landmarken von der gegenüberliegenden Seite zu integrieren. Darum ist aus Effizienzgründen anzunehmen, daß Information über Objekte vom Betrachter frühzeitig als aktuell unwichtig ausgefiltert werden, die mit den aktuellen Intentionen nicht konform sind.

#### 4.1.4 Das Raumintegrationsmodul

Um sich in einer Situation effektiv und effizient verhalten zu können, benötigt der Agent integrierte räumliche Repräsentationen von Referenzsystemen, visuell ausgewählten Objekten und ausgezeichneten räumlichen Relationen. Mittels Referenzsystemen läßt sich die Situation räumlich strukturieren und Objekte durch räumliche Relationen zueinander und zum Akteur in Beziehung setzen. In den vorigen Kapiteln ist beschrieben worden, wie das visuelle Objektauswahl- und das Pfadsuchmodul visuell saliente Landmarken bzw. Pfadelemente bestimmen, die in einer gegebenen Situation für den Agenten zur Erfüllung seiner Aufgabe relevant sind. Die Funktionen des *Raumintegrationsmoduls* integrieren diese Entitäten zu einer räumlichen Repräsentation, der sogenannten *räumlichen Konfigurationsbeschreibung (RKB)*. Räumliche Beziehungen zwischen Entitäten werden durch räumliche Relationen hergestellt. Es wird dazu zwischen *topologischen* und *projektiven* (deiktisch und intrinsisch) räumlichen Relationen unterschieden. Der funktionale Aufbau des Raumintegrationsmoduls ist in Abbildung 4.4 dargestellt.

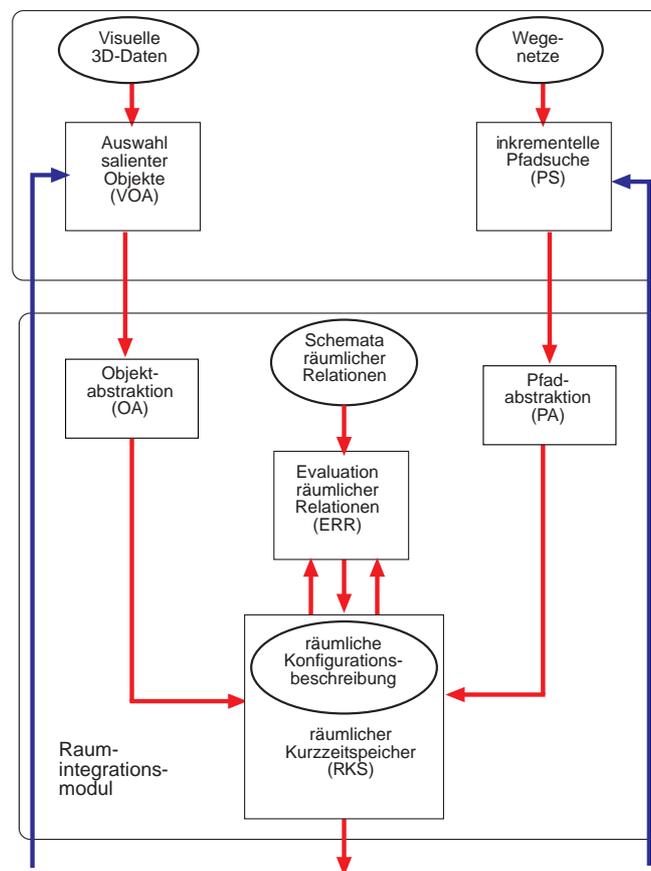


Abbildung 4.4: Der funktionale Aufbau des Raumintegrationsmoduls

Im weiteren Verlauf werden die verschiedenen Referenzsysteme skizziert, die die Basis für die Etablierung räumlichen Wissens bilden. Anschließend werden die unterschiedlichen Arten von räumlichen Relationen und die Bedeutung des räumlichen Kurzzeitspeichers vorgestellt.

##### 4.1.4.1 Räumliche Referenzsysteme

Räumliche Referenzsysteme lassen sich danach unterscheiden, ob sie *egozentrisch*, und damit innerhalb, oder *allozentrisch* außerhalb des Agenten verankert sind (vgl. 2.3.1). Übliche Defi-

nitionen von allozentrischen Referenzsystemen abstrahieren derart vom Betrachterstandpunkt des Agenten, daß sie für einen sich bewegenden Agenten nicht direkt übernommen werden können (vgl. Kapitel 2.3.1). Denn die Eigenbewegung des Agenten bewirkt gleichsam eine dynamische Veränderung der wahrgenommenen, räumlichen Situation. In statischen Situationen (vgl. [Her86], [Gap94]) kann vom Betrachterstandpunkt abstrahiert werden. In dynamischen Situationen ist jedoch die Lokation des Agenten, und damit verbunden vor allem das egozentrische Referenzsystem, im Mittelpunkt der Analyse. Die Orientierung und Verwendung von Referenzsystemen ändert sich dabei dynamisch. Um die Bewegung und die Perspektive getrennt voneinander berücksichtigen zu können, wird das egozentrische Referenzsystem in ein *primäres* und ein *sekundäres* Referenzsystem unterteilt. Ein *primäres Referenzsystem* ist durch die vom Agenten aktuell ausgeführte Hauptrichtung der Bewegung, der Handlung oder Perzeption festgelegt. Daneben ist ein Agent durch kurzfristige Umorientierung in der Lage, räumliche Bereiche außerhalb der Hauptrichtung zu fokussieren, wozu er ein egozentrisches, *sekundäres Referenzsystem*, bei gleichzeitigem Erhalt des primären Referenzsystems verwendet.<sup>10</sup> Um Objekte, bezogen auf die Lokation des Agenten, zueinander in Beziehung zu setzen, wird entsprechend dem Prinzip des geringsten kognitiven Aufwands ein geeignetes Referenzsystem ausgewählt.<sup>11</sup> Die Verwendung eines sekundären sowie des nachfolgend beschriebenen virtuellen Referenzsystems kann auch sprachlich explizit angegeben werden, wie zum Beispiel durch folgende Anweisung: „Wenn Sie nach rechts schauen, dann . . .“

Die Evaluation von räumlichen Relationen zwischen zwei Objekten beruht auf der Etablierung einer räumlichen Lagebeziehung zwischen einem Referenzobjekt (refo) und einem Lokalisationsobjekt (lo). Dazu wird bezüglich der Lokation des Referenzobjekts ein allozentrisches, *virtuelles Referenzsystem* generiert, welches in einer Lokation im Raum verankert und durch die Hauptachsen des Referenzobjektes vorgegeben ist. Ein virtuelles Referenzsystem wird aus einem primären oder sekundären Referenzsystem durch Translationen und Rotationen abgeleitet.<sup>12</sup> Unter der Annahme, daß Translationen sehr effizient ausgeführt werden können, ist die Minimierung von mentalen Rotationen beim Aufbau von virtuellen Referenzsystemen das wesentlich Designkriterium. Eine formale Vorstellung dieser Referenzsysteme erfolgt in Kapitel 4.2.4.1.

#### 4.1.4.2 Evaluation räumlicher Relationen

Ist bezüglich einer Situation ein geeignetes Referenzsystem ausgewählt worden, werden zwischen ausgewählten Objekten und dem Agenten räumliche Relationen etabliert. Grundlage für Raumbeschreibungen ist eine ausreichende raumbezogene Repräsentation, um geeignete raumsprachliche Ausdrücke zu instantiieren. Allgemein repräsentieren verschiedene Typen von Relationen unterschiedliche Arten von konzeptuellen Beziehungen (vgl. Kapitel 2.3.2). In dieser Arbeit werden drei Typen von räumlichen Relationen verwendet, die als Basis der Raumrepräsentationen verwendet werden. *Topologische räumliche Relationen* repräsentieren Beziehungen, die referenzsystemunabhängig sind. Sie explizieren grobe physische Zusammenhänge auf qualitative Weise. *Deiktische* und *intrinsische räumliche Relationen* hingegen hängen vom verwendeten Re-

<sup>10</sup>Zugrunde liegt die Hypothese, daß der Sprecher annimmt, daß der Rezipient räumliche Referenzsysteme in derselben Weise wie der Sprecher etabliert. Es ist zu vermuten, daß eine solche Strategie erlernt werden muß, und daß sie demzufolge sprachabhängig ist (vgl. [GM95], [GW95]).

<sup>11</sup>Wie empirische Untersuchungen zeigen, tendieren Probanden dazu, den kognitiven Aufwand des Rezipienten zu minimieren. Da im Falle der inkrementellen Wegbeschreibungen, so wie sie hier verstanden werden, davon ausgegangen wird, daß der Sprecher und der Rezipient sich nebeneinander in die gleiche Richtung bewegen, wird in dieser Arbeit der Aspekt der Einnahme einer Rezipientenperspektive nicht weiter diskutiert. Generell ist eine solche Einschränkung für inkrementelle Wegbeschreibungen nicht notwendig. Jedoch ist eine Beschreibung für einen Rezipienten am einfachsten zu verstehen, wenn sie bezüglich seines egozentrischen Referenzsystems gegeben werden (vgl. [HGH91]).

<sup>12</sup>Im Sinne von Herrmanns „Sichhineinversetzen“ (vgl. [HGH91]).

ferenzsystem ab. Topologische, deiktische und intrinsische räumliche Relationen repräsentieren *sprachunabhängige* Beziehungen zwischen Objekten. Erst im Sprachgenerierungsmodul werden diese Relationen auf Lexeme, vor allem auf räumliche Präpositionen, abgebildet (vgl. Abbildung 4.5).

Deiktische Relationen repräsentieren explizit, wie sich die räumliche Lage eines Objektes zur Lokation des Agenten verhält. Die Evaluation von deiktischen Relationen wird bezüglich eines primären oder sekundären Referenzsystems durchgeführt. Intrinsische Relationen stellen Beziehungen zwischen einem Referenzobjekt, ungleich dem Agenten, und einem Lokalisationsobjekt her. Wie empirische Befunde belegen (siehe Kapitel 2.3.1), stellt die Evaluation von intrinsischen Relationen höhere Anforderungen an kognitive Kapazitäten als deiktische Relationen. Eine effizientere Evaluation wird durch eine minimal-aufwendige Etablierung eines virtuellen Referenzsystem erreicht. Durch Antizipation des Kriteriums des minimalen Aufwands bestimmt der Rezipient, welches Referenzsystem der Raumbeschreibung der Sprecher zugrunde gelegt hat. Besondere Verwendung findet der Zusammenhang zwischen der Auswahl geeigneter Referenzsysteme und räumlichen Relationen in dynamischen Situationen, in denen die räumlichen Konfigurationen an die Veränderungen der perzipierten Situation angepaßt werden.

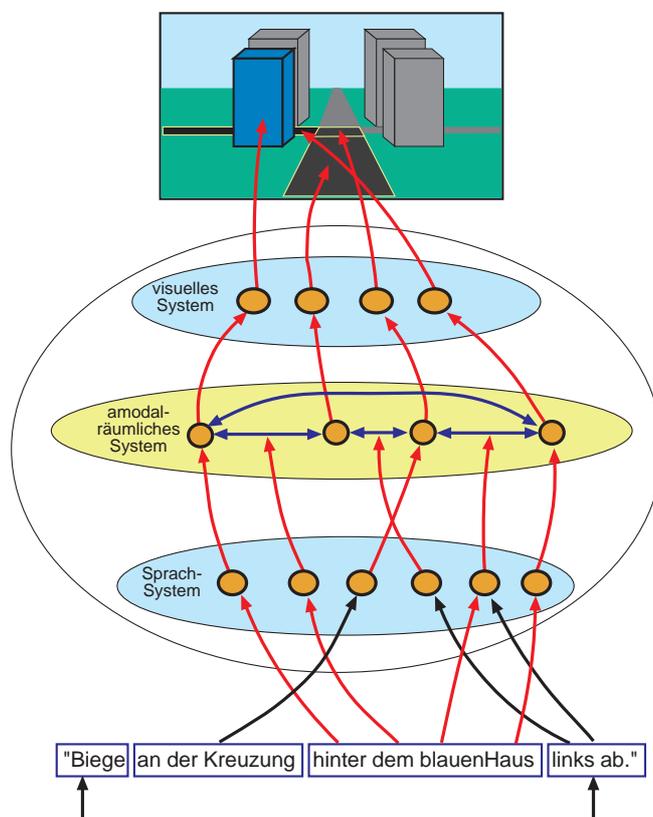


Abbildung 4.5: Der konzeptuelle Zusammenhang zwischen den Repräsentationsebenen

Um den Unterschied zwischen der *amodal-räumlichen* und *sprachlichen Ebene* zu kennzeichnen, werden räumliche Relationen auf der konzeptuellen Ebene durch englischsprachige Ausdrücke beschrieben und durch " #" prä- und postfigiert, wie zum Beispiel bei #right-of#. Dies kennzeichnet eine explizite Unterscheidung zwischen räumlichen Relationen und räumlichen Präpositionen bzw. deren Lexemen ( siehe Abbildung 4.5). Die restringierte Menge dieser Relationen besteht aus den folgenden Elementen, die allgemein der Menge der projektiven Relationen zugeordnet wird (vgl. [RS88], [Gap94]): #left-of#, #right-of#, in-front-of#, #behind#, #above#

und #under#. Diese Relationen korrespondieren in gebietskonstituierender Weise zu kartesischen Referenzsystemen in drei Dimensionen (vgl. Kapitel 2.3.1). Inhärent ist diesen Relationen, daß sie einen weiteren Filter für das repräsentierte räumliche Wissen festlegen. Information über quantitative Lagebeziehungen von Objekten zueinander wird durch eine qualitative Repräsentation ersetzt. Relationen dieser Art sind nicht durch zweiwertige, logische Prädikate, sondern über reellwertige Anwendbarkeitsgrade definiert.<sup>13</sup> Insgesamt wird die räumliche Repräsentation, die durch eine räumliche Konfigurationsbeschreibung etabliert wird, kompakter, was nicht zuletzt in Bezug auf die Effizienz assoziierter Funktionen Auswirkungen hat. Eine Diskussion der Evaluation räumlicher Relationen erfolgt in Kapitel 4.2.4.2.

#### 4.1.4.3 Räumlicher Kurzzeitspeicher

Der räumliche Kurzzeitspeicher integriert Funktionen zur Etablierung von Referenzsystemen und räumlichen Relationen, sowie den Aufbau und die Verwaltung von räumlichen Konfigurationsbeschreibungen. Entsprechend den Erkenntnissen aus den Bereichen der kognitiven Psychologie, wird räumliches Wissen auch langfristig gespeichert, was ein persistentes Speichermedium und entsprechende Repräsentationen und Funktionen voraussetzt. Engelkamp und Zimmer beispielsweise gehen davon aus, daß ein solches *räumliches Langzeitgedächtnis* (vgl. [EZ94]) räumliches Wissen mit einem räumlichen Kurzzeitgedächtnis austauscht. Demzufolge ist der räumliche Kurzzeitspeicher, so wie er in dieser Arbeit verwendet wird, nur ein Teil des Raumintegrationsmoduls. Weiterhin ist dieser Kurzzeitspeicher auf relationales Wissen eingeschränkt. Entsprechend einer "dual-coding"-Theorie ist zu fordern, daß auch depiktionale Repräsentationen berücksichtigt werden (vgl. [HP88], [Pri93]). Unter der Annahme, daß für eine inkrementelle Beschreibung eines Weges kurzzeites, relationales Wissen ausreichend ist, wird entsprechend nur ein eingeschränkter Kurzzeitspeicher berücksichtigt<sup>14</sup>. Die funktionale Struktur des räumlichen Kurzzeitspeichers wird in Kapitel 4.2.4.3 vorgestellt.

#### 4.1.5 Das Navigationsmodul

Neben der perzeptionsgebundenen Informationsaufnahme ist in das Modell weiterhin ein Navigationsmodul integriert. Dies erlaubt dem Agenten eine qualitative Ansteuerung von Bewegungsaktionen. Es wird sich in dieser Arbeit auf eine abstrakte Beschreibung der Navigation beschränkt, die sich an Kuipers Ebene des *sensomotorischen Wissens* anlehnt (vgl. [Kui78]). Durch die beiden Aktionen *MOVE* und *ROTATE* ist der Agent in der Lage, sich in einer Situation *translatorisch* oder *rotational* zu bewegen. Nach Beendigung einer Aktion wird das primäre Referenzsystem an die neue Ausrichtung und Lokation des Agenten angepaßt. Bei einer Translation wird die neue Lokation und Orientierung der Bewegungsrichtung in diskreten Abständen an das Raumintegrationsmodul übermittelt, welches daraus bestimmt, ob Referenzsysteme und das aktuelle raumbezogene Wissen angepaßt werden müssen.

Es ist prinzipiell notwendig, das raumbezogene Wissen über einen Pfadabschnitt mit Information, die aus der Situation extrahiert worden ist, zu vergleichen. Da es sich im vorliegenden Modell nicht um eine physikalische Realisierung handelt, wird eine dediziertere Instantiierung vernachlässigt. Bei einer Integration des Modells in einen physikalischen Agenten bedarf es einer genaueren Formulierung dieser Instantiierung. Eine vollständige Modellierung dieses Moduls fordert die Integration kontrolltheoretischer Ansätze zu physischen Agenten (z.B. [Bro86]), was aber außerhalb des Rahmens dieser Arbeit liegt.

<sup>13</sup>Es bestehen Ähnlichkeiten des Konzepts der Anwendbarkeitsgrade zu Konzepten der "fuzzy set theory" (z.B. [Zad65], [Lak73]).

<sup>14</sup>Für einen ersten Ansatz zur Modellierung eines vollständigen Raumintegrationsmoduls siehe [GKS89].

### 4.1.6 Das Sprachproduktionsmodul

Inkrementelle Wegbeschreibungen sind in ihrer linguistischen Struktur schematisch aufgebaut (vgl. Kapitel 3). Obschon auf der Ebene der Oberflächenstrukturen unterschiedliche Ausprägungen von Beschreibungen einer Situation generiert werden können, sind diese durch eine kleine Menge von inhaltlichen Strukturen und grammatikalischen Regeln determiniert (vgl. Kapitel 3). Gestützt durch die hier vorgestellten Untersuchungen und durch die Ergebnisse von Arbeiten zu vollständigen Wegbeschreibungen (vgl. Kapitel 2.7), ist die Generierung von Wegbeschreibungen semantisch, syntaktisch sowie kommunikativ stark strukturiert. Die Inhaltsstrukturen, die einer Beschreibung zugrunde liegen, bestehen aus einer begrenzten Menge an strukturellen Elementen, die entsprechend situationsbedingter Beschränkungen miteinander kombiniert werden. Oberflächenstrukturen inkrementeller Wegbeschreibungen werden konstruktiv aus Inhaltsstrukturen durch Anwendung einer beschränkten Menge von grammatikalischen Regeln generiert.

Das Ziel des Agenten ist es, eine Wegbeschreibung bezüglich einer Situation derart zu generieren, daß diese einerseits in der zur Verfügung stehenden Zeit gegeben und andererseits die Situation und die entsprechenden Aktionen eindeutig identifiziert werden kann. Werden beide Bedingungen erfüllt, so wird eine inkrementelle Wegbeschreibung als *adäquat* bezeichnet. Inkrementelle Wegbeschreibungen bestehen i.a. aus fünf *Inhaltskategorien*, die zeitliches, räumliches und intentionales Wissen explizieren (vgl. Kapitel 3). Über Generierungsfunktionen werden bezüglich einer Situation geeignete Inhaltskategorien ausgewählt und instantiiert.

Entsprechend den empirischen Untersuchungen lassen sich vier Inhaltskategorien unterscheiden (vgl. Kapitel 3.2): *WANN-K*, *WAS-K*, *WOHIN-K*, *WO-K*. Aus Gründen der Vollständigkeit wird noch zusätzlich eine *WER-K* Kategorie angenommen, die implizit linguistische Eigenschaften an andere Elemente vererbt. Die *WER-K* Kategorie wird durch eine Referenz auf den Rezipienten instantiiert.

Die formale Beschreibung unterteilt sich in den Schritt der Generierung und der Präsentation. In der Generierungsphase wird eine geeignete semantische Repräsentation der jeweiligen Aussage und eine sprachliche Transformation dieser determiniert. Jedoch wird entsprechend den empirischen Befunden eine Beschreibung erst dann präsentiert, wenn eine geeignete raum-zeitliche Distanz zum nächsten Verzweigungspunkt erreicht worden ist (vgl. Kapitel 3.2). Eine Beschreibung der sprachbezogenen Funktionen zur Generierung von Oberflächenstrukturen ist in Kapitel 4.2.6 gegeben. Im folgenden werden die einzelnen Funktionen der Module explizit vorgestellt.

## 4.2 Funktionale Beschreibung des kognitiven Wegbeschreibungsagenten

Der Agent agiert in einem dreidimensionalen Modell des Campus der Universität des Saarlandes, welcher mit *Landmarken*, *Pfaden* und *Verzweigungspunkten* drei wesentliche Elemente einer stadtähnlichen Umgebung besitzt (vgl. Abbildung 4.6 und Kapitel 2.3). Mit diesen Elementen korrespondieren *Straßen*, *Kreuzungen*, *Plätze*, *Kreisverkehre*, *Gebäude* und *Grundflächen*.<sup>15</sup> Die Umgebung repräsentiert somit eine Granularitätsstufe, durch die ein Betrachter Zugang zu einer typischen, alltäglichen Umgebung findet.

In einer Situation dieser dreidimensionalen Umgebung seligiert, internalisiert, verarbeitet und verwendet der Agent visuelle Information als Basis für sein sprachliches und navigationsbezogenes Verhalten. Es wird im folgenden das funktionale Modell eines kognitiven Agenten erläutert, welcher in der Lage ist, sich in einer dreidimensionalen Umgebung zu bewegen und adäquate inkrementelle Wegbeschreibungen zu generieren. Grob lassen sich Funktionen unterscheiden, die

<sup>15</sup>Nicht modelliert sind bisher dynamische Objekte, wie beispielsweise Autos, Fahrräder oder andere Personen.

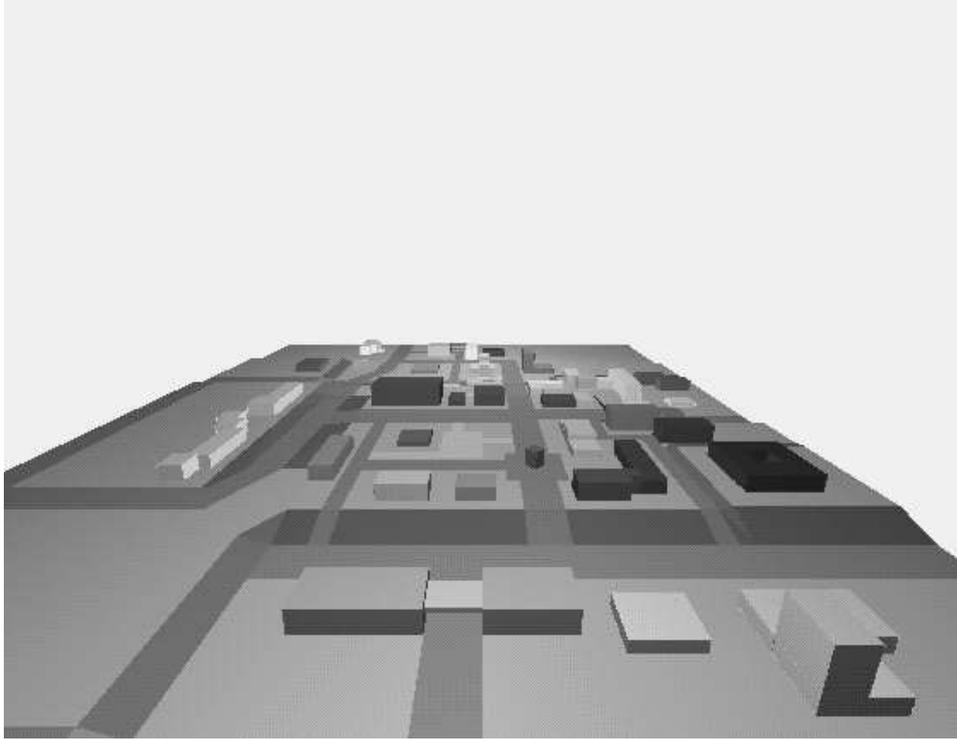


Abbildung 4.6: Dreidimensionale Umgebung des Campus der Universität des Saarlandes

in der Schnittstelle zwischen der Umgebung und dem Agenten liegen, und zum anderen solche, die davon abgekapselt und von außen nicht direkt zugänglich sind. Das äußerliche Verhalten des Agenten besteht in seinen Bewegungen in der Umgebung und seinen verbalen Beschreibungen, die er einem ihn begleitenden Rezipienten vermittelt. Der Agent ist maximal kooperativ und antizipiert in eingeschränktem Maße die Intentionen und das Wissen über den Adressaten, um sein navigationsbezogenes und verbales Verhalten darauf einzustellen.

Die Funktionen des Steuerungsmoduls *K-M* (*globale Koordination* und *Evaluation temporaler Beschränkungen*) steuern die *Aktivierung* und *Deaktivierung* der anderen Module. Die Objektauswahl *VOA-M* etabliert eine Referenzbeziehung zwischen perzipierten Landmarken und räumliches Wissen (*Auswahl salienter Objekte*). Das Pfadsuchmodul determiniert heuristisch und inkrementell konsekutive Folgen von Pfadabschnitten. Referenzierte Objekte und Pfadabschnitte werden im Raumintegrationsmodul *RI-M* zu multiplen Repräsentationen abstrahiert (*Objektabstraktion* und *Pfadabstraktion*), zueinander und zum Betrachter in Beziehung gesetzt (*Evaluation räumlicher Relationen*) und letztendlich in die aktuelle *räumliche Konfigurationsbeschreibung* integriert (*räumlicher Kurzzeitspeicher*). Über Wissen des räumlichen Kurzzeitspeichers erhalten das Objektauswahlmodul und das Pfadsuchmodul Angaben zur Parametrisierung relevanter Fokussierungen. Das Navigationsmodul *NV-M* erlaubt es dem Agenten, sich translatorisch und rotational zu bewegen (*Navigation*) und die aktuelle Ausrichtung relativ zur Situation zu bestimmen (*Orientierung*). Das Raumintegrationsmodul steuert einerseits das Navigationsmodul zwecks Fortbewegung in der 3D-Umgebung, erhält aber andererseits raumbezogene Information über die aktuelle Lokation und Orientierung zurück, woran Referenzsysteme angepaßt werden. Mittels raumbezogener, zeitbezogener und intentionaler Information wird im Sprachmodul der semantische Gehalt einer Wegbeschreibung bestimmt (*Auswahl geeigneter Inhaltsstrukturen*), in Oberflächenstrukturen transformiert und mittels sprachgenerativer Funktionen in eine natürlich-sprachliche Beschreibung umgesetzt. Die Aktivierung der Präsentation einer Beschreibung ist

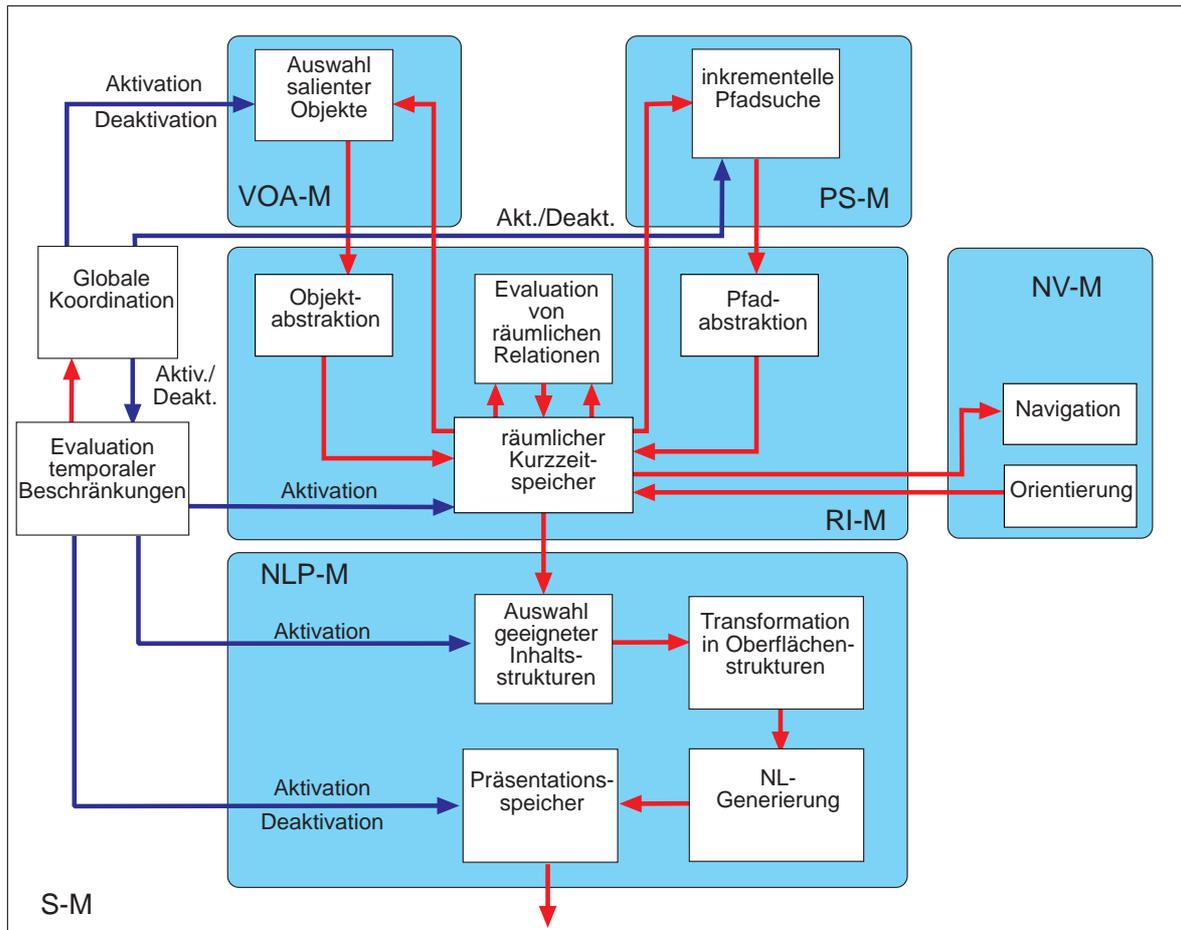


Abbildung 4.7: Das vollständige formale Modell des kognitiven Wegbeschreibungsagenten

von raum-zeitlichen Beschränkungen abhängig und bedingt die Verwendung von “anytime”-Mechanismen.

In den folgenden Kapiteln werden die Funktionen der einzelnen Module beschrieben.

#### 4.2.1 Funktionale Beschreibung des Steuerungsmoduls

Die Aufgabe des Steuerungsmoduls ist die Bestimmung von *Ablaufplänen* (schedules), durch die Phasen aktiviert und deaktiviert werden. Über Phasen wiederum werden Module angesteuert, die wiederum die Evaluation von Funktionen anstoßen. Das Zusammenspiel von Modulen und Funktionen wird durch zeitliche Restriktionen eingeschränkt, die durch die Eigenbewegung des Agenten und die Umgebung gegeben sind.

Im weiteren Verlauf wird diskutiert, wie zeitliche Restriktionen bei der Generierung berücksichtigt werden. Dazu wird die Zeit vom Agenten geschätzt, über die er für eine Beschreibung verfügt. Dieser Zeitrahmen wird auf die einzelnen Module verteilt. Anschließend wird vorgestellt, welche Phasen durch das Steuerungsmodul angesteuert werden können und wie diese die einzelnen Verarbeitungsschritte determinieren.

#### 4.2.1.1 Zeitrahmen und Zeitbeschränkungen

Wie die empirischen Untersuchungen zeigen (vgl. Kapitel 3), lassen sich inkrementelle<sup>16</sup> Wegbeschreibungen in fünf Klassen  $T_1, \dots, T_5$  einteilen, die in direkter Abhängigkeit vom zeitlichen Rahmen der Situation stehen (siehe Tabelle 3.1 in Kapitel 3.2).

Besteht die Beschreibung einer Situation und einer Aktion aus mehreren Teilbeschreibungen, so lassen sich Abhängigkeiten zwischen diesen durch Zuordnung zu Zeitrahmen formalisieren. Der Agent versucht dabei, den Rezipienten nicht mit Information zu überfordern, sondern reduziert die Komplexität auf eine Untermenge von möglichen Beschreibungen. Wird eine Teilbeschreibung zu einem Zeitrahmen  $T_i$  gegeben, so wird frühestens zu einem Zeitrahmen  $T_{i+2}$  eine weitere Teilbeschreibung präsentiert. Ein solches Verhalten ist nur dann möglich, wenn  $i \in \{2, 3\}$  gilt, da während  $T_1$  ausschließlich unterstützende Beschreibungen generiert werden (vgl. Kapitel 3.2). Dadurch ergeben sich empirisch belegte Abhängigkeiten, wie sie in Abbildung 4.8 dargestellt sind.

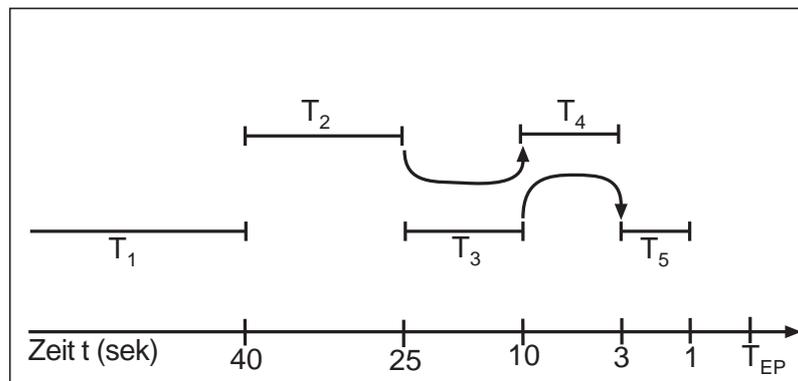


Abbildung 4.8: Zeitintervalle, die den inkrementellen Wegbeschreibungen prototypisch zugeordnet sind.

Um bestimmen zu können, in welchem Zeitrahmen sich der Agent zu einem gegebenen Zeitpunkt befindet, wird der Zeitrahmen  $T_{guess}$  vom Agenten geschätzt. Dieser Zeitrahmen determiniert, wieviel Zeit dem Agenten bis zum Erreichen des nächsten Verzweigungspunktes und für die gesamte kommunikative Handlung zur Verfügung steht. Dazu wird  $T_{guess}$  in vier Teilzeitrahmen aufgeteilt:

1.  $T_{proc}$ : Zeitrahmen zur Generierung einer inkrementellen Wegbeschreibung (IWB)
2.  $T_{pres}$ : Zeitrahmen zur Präsentation einer IWB

<sup>16</sup>Finkler schlägt eine Unterscheidung in *quantitative* und *qualitative Inkrementalität* vor (vgl. [Fin96]). Ein Prozeß ist demnach quantitativ inkrementell, wenn die Ausgabe in seiner Gesamtheit erst dann zu erkennen ist, wenn alle Einzelschrittausgaben vollständig gegeben worden sind. Jede Einzelausgabe ist in sich vollständig. Im Gegensatz dazu ist ein Prozeß qualitativ inkrementell, wenn jede Einzelausgabe die Qualität der gesamten Ausgabe verbessert. Ein Beispiel für eine quantitative Inkrementalität ist der zeilenweise Bildaufbau beim Fernseher. Im Gegensatz dazu wird bei einem Mikroskop ein unscharfes Bild qualitativ inkrementell aufgebaut, wenn man die Schärfe verbessert. Diese Unterscheidung wird durch den hier vorgestellten Ansatz "überlappt, da zuerst eine minimale Beschreibung generiert wird, die die Situation vollständig beschreibt. Danach bzw. nebenläufig dazu wird, falls zeitliche Restriktionen dies erlauben, eine Erweiterung dieser Beschreibung generiert, die ihrerseits wiederum vollständig ist. Eine rein qualitative Inkrementalität kann in einfacher Weise dadurch erreicht werden, daß die Erweiterung nicht mehr durch eine eigene vollständige Beschreibung verbalisiert wird, sondern durch Anhängung der neuen Information an die vorhergehende Beschreibung. Da ein solches sprachliches Verhalten im Korpus nicht vorkam, wird dies im folgenden nicht weiter betrachtet.

3.  $T_{listener}$ : Angenommener Zeitrahmen, den der Rezipient zum Verstehen einer IWB benötigt
4.  $T_{action}$ : Geschätzter Zeitrahmen zur Ausführungsaktivierung<sup>17</sup>

Die beiden Zeitrahmen  $T_{pres}$  und  $T_{listener}$  hängen voneinander ab, da die Komplexität der Beschreibung approximativ durch die Länge der IWB gegeben ist.<sup>18</sup> Es wird angenommen, daß sich mit einer längeren Beschreibung auch die Verstehenszeit des Rezipienten verlängert. Ein zeitliche Beschränkung, die im Falle der inkrementellen Wegbeschreibungen nicht überschritten werden darf, legt die folgende Ungleichung fest:

$$T_{guess} \geq T_{proc} + T_{pres} + T_{listener} + T_{action}$$

Durch Modifikation der internen Einstellungen des Agenten bezüglich den antizipieren Eigenschaften des Rezipienten, kann die gesamte Generierung beeinflusst werden. Geht der Agent davon aus, daß der Hörer eine lange Zeitspanne benötigt, um eine Beschreibung zu verstehen, so bleibt weniger Zeit für die Generierung zur Verfügung, so daß die Qualität der Beschreibung abnimmt. Ergo expliziert  $T_{guess}$  die zeitliche Ressourcenlimitation des Agenten zur Beschreibung des Weges (siehe Kapitel 4.2.1).<sup>19</sup> Nur wenn diese Beschränkung erfüllt sind, ist sichergestellt, daß die Beschreibung zum richtigen Zeitpunkt erfolgt.<sup>20</sup>

Die Verbindung zwischen den Zeitrahmen  $T_i$  ( $i = 1 \dots 5$ ) und  $T_{guess}$  erfolgt über  $T_{pres}$ . Der Agent bestimmt, zu welchem Zeitrahmen  $T_i$  er intendiert, eine Beschreibung dem Rezipienten präsentierbar zur Verfügung zu stellen. Dafür ist die Optimierungsaufgabe zu lösen, die nach dem Zeitrahmen  $T_i$  mit minimalem  $i$  sucht, welcher die untere Grenze der Addition der beiden Zeitrahmen  $T_{proc}$  und  $T_{pres}$  enthält:

$$\text{Suche minimales } i: \text{lb}^{21}(T_{proc} + T_{pres}) \in T_i$$

Je weniger komplex die Generierung einer Beschreibung ist, um so kürzer sind  $T_{proc}$  und  $T_{pres}$ . Da der Agent aber maximal kooperativ ist, sucht er ein  $T_i$ , welches eine möglichst detaillierte Beschreibung ermöglicht. Die Länge von  $T_{proc}$  und  $T_{pres}$  sind von der Komplexität

<sup>17</sup>Bei Fußgängern kann dieser Zeitrahmen vernachlässigt werden. Bei der Navigation von größeren Objekten, wie beispielsweise Öltankschiffen, ist dies jedoch ein ganz entscheidender Faktor, der ein Vielfaches der Zeit einnimmt, die man für die Generierung einer Beschreibung benötigt.

<sup>18</sup>Allgemein ist die Komplexität eines Sates durch die Phrasenstruktur bestimmt (vgl. [Kin77]). Im Falle inkrementeller Wegbeschreibungen hingegen werden durchweg einfache Satzkonstruktionen verwendet, bei denen eine einzelne Phrase auf der Oberflächenstrukturebene fast ausschließlich in maximal drei Worte transformiert wird. Bei kurzen Beschreibungen werden Phrasen meist durch einzelne Worte verbalisiert. Somit erscheint die approximative Abschätzung der Komplexität des Inhalts einer Beschreibung durch die Anzahl der Einzelworte im Falle inkrementeller Wegbeschreibungen gerechtfertigt.

<sup>19</sup>Ein solches Verhalten steht im Einklang mit der Hypothese, daß ein menschlicher Agent über beschränkte Ressourcen verfügt und dabei mit steigender Verfügung einer Ressource auch die Qualität des Ergebnisses, im vorliegenden Falle des Sprachverhaltens, steigert (vgl. auch [Cha87], [DB87], [RW91]).

<sup>20</sup>Da auch die Probanden nicht in jedem Fall zum richtigen Zeitpunkt eine Beschreibung geben konnten, kann davon ausgegangen werden, daß auch sie nicht immer die Situation zeitlich korrekt einzuschätzen wissen. Dies kann zum einen daher rühren, daß sie die Situation falsch eingeschätzt haben, oder aber bereits von Anfang an wußten, daß sie die Aufgabe nicht erfüllen konnten. Im letzteren Fall waren sie dann nicht in der Lage, die Beschreibung geeignet zu verkürzen (vgl. Kapitel 3).

<sup>21</sup>Die Abkürzung *lb* steht für die untere Grenze (lower bound).

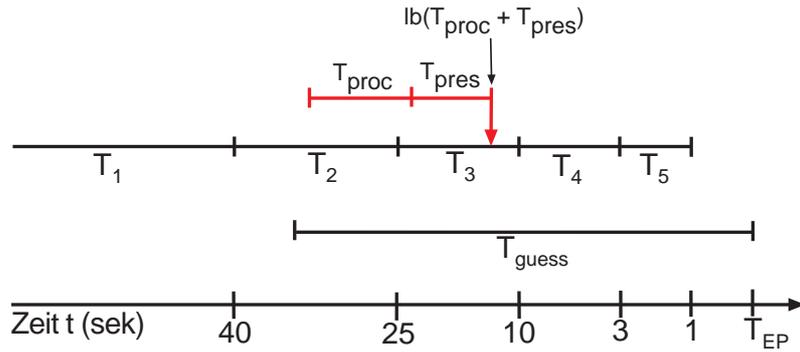


Abbildung 4.9: Zeitliches Optimierungsproblem

der Beschreibung abhängig, welche wiederum einem bestimmten  $T_i$  zugeordnet ist. Dadurch sind  $T_{proc}$  und  $T_{pres}$  direkt von  $T_i$  abhängig. Somit kann obige Bedingung wie folgt umgeformt werden:

$$\text{Suche minimales } i: lb(T_{proc}(T_i) + T_{pres}(T_i)) \geq lb(T_i).$$

Aus den Untersuchungen konnten Mittelwerte für  $T_{proc}(T_i)$  und  $T_{pres}(T_i)$  abgeleitet werden, so daß diese Suche nach einem minimalen  $i$  effizient evaluiert werden kann (ein Beispiel hierzu ist in Abbildung 4.9<sup>22</sup> dargestellt).

Im folgenden wird vorgestellt, wie im Modell der Zeiträume  $T_{proc}$  dekomponiert und dadurch das Verhalten des gesamten Modells bestimmt wird.

#### 4.2.1.2 Dekomposition des Zeiträhmens $T_{proc}$

Für das sprachliche Verhalten des Agenten ist in erster Linie der Zeiträhmen  $T_{proc}$  verantwortlich. Die Funktionen zur *Evaluation temporaler Beschränkungen* dekomponieren  $T_{proc}$  in kleinere Zeiträhmen. Eine statische Dekomposition von  $T_{proc}$  auf Module schränkt die Adaptivität des Agenten erheblich ein, was eine geringe Anpassungsfähigkeit zur Folge hat. Adaptiveres Verhalten erlaubt eine Dekomposition in Abhängigkeit vom vorgegebenen Zeiträhmen  $T_{proc}$ . In diesen, den  $T_i$  assoziierten Kombinationen von Inhaltsstrukturen, ist der Zeiträhmen  $T_{proc}(T_i)$ <sup>23</sup> unterschiedlich lang, was Auswirkungen auf die Qualität der Beschreibungen hat. Die empirisch belegte Einteilung in fünf Zeiträhmen  $T_1$  bis  $T_5$  erlaubt eine Diskretisierung des Problemraums, so daß vermeiden wird, permanent den Wert von  $T_{proc}$  überprüfen und berücksichtigen zu müssen. Wie die empirischen Untersuchungen zeigen (vgl. Kapitel 3), reicht eine verhältnismäßig grobe Schätzung von  $T_{proc}$  aus, um eine adäquate Beschreibung zu generieren. Dieser Umstand macht das Verhalten des Agenten robuster gegenüber kleinen Änderungen in der Umgebung und gegenüber Fehleinschätzungen.

Unter der Annahme, daß der Sprecher versucht,  $T_{proc}$  maximal auszuschöpfen, ergibt sich folgende Ungleichung (vgl. auch Abbildung 4.10):

<sup>22</sup>Aus Platzgründen sind die eingezeichneten Zeiträhmen für  $T_{proc}$  und  $T_{pres}$  nicht in Sekunden skaliert eingezeichnet.

<sup>23</sup>In Fällen, in denen es nicht von primärer Bedeutung ist oder aus dem Kontext hervorgeht, bezüglich welchem  $T_i$  eine Beschreibung generiert wird, wird die  $T_{proc}(T_i)$  durch  $T_{proc}$  abgekürzt.

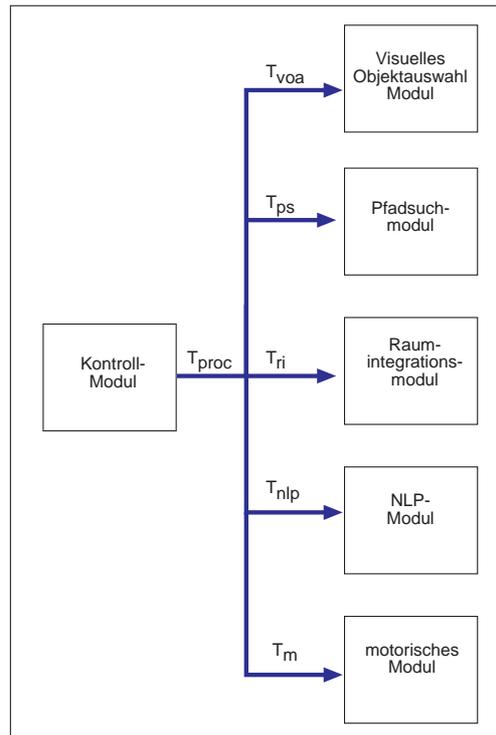


Abbildung 4.10: Kontrollfluß bedingt durch die zeitliche Beschränkung  $T_{proc}$

$$T_{proc} \geq T_{object} + T_{ps} + T_{ri} + T_{nlp} + T_{nv}.$$

Die Funktionen der einzelnen Module werden nicht in beliebiger, sondern in partiell koordinierter Weise aktiviert. Zwischen Funktionen bestehen *Sequentialisierungsrelationen*, die angeben, ob zwei Funktionen nur strikt sequentiell oder auch nebenläufig evaluiert werden können.

Durch Abhängigkeiten zwischen Modulen wird angegeben, welche Module zueinander *nebenläufig* beziehungsweise *sequentiell* evaluiert werden können. Ein Modul ist zu einem anderen *nebenläufig*, wenn es einen Zeitpunkt  $t$  gibt, an dem beide Module gleichzeitig aktiviert sind. Zwei Module stehen zueinander in sequentieller Beziehung, wenn sie nicht nebenläufig sind.

Zwischen den eingeführten Modulen des Agenten gibt es *Sequentialisierungsrelationen*, die in Tabelle 4.1 abgebildet sind. Hierbei drückt eine Sequentialisierungsabhängigkeit  $s(x, y)$ , aus, daß ein Modul  $x$  nicht gleichzeitig mit Modul  $y$  aktiv sein kann. Diese beiden Module können zeitlich voneinander getrennt aktiv sein, hingegen muß zu einem Zeitpunkt mindestens eines von beiden de-aktiviert sein.

Ist beispielsweise das visuelle Objektauswahlmodul  $VOA-M$  aktiv, kann das Pfadsuchmodul  $PS-M$  nicht aktiv sein und vice versa (vgl. Tabelle 4.1). Dies repräsentiert, daß der Agent nicht gleichzeitig auf einer Karte lesen und die visuelle Information der Situation analysieren kann. Im Raumintegrationsmodul  $RI-M$  wird dies durch zwei Verarbeitungswege reflektiert. Entweder kann  $RI-M$  mit  $VOA-M$  oder mit  $PS-M$  sequentiell interagieren (vgl. Abbildung 4.11, 4.12). Eine weitere Sequentialisierungsabhängigkeit besteht zwischen  $RI-M$  und dem Sprachgenerierungsmodul  $NLP-M$ . Da  $NLP-M$  auf Strukturen zurückgreift, die durch  $RI-M$  generiert werden, wird  $NLP-M$  solange deaktiviert, wie  $RI-M$  aktiv ist. Ansonsten kommt es zu Überschreibungseffekten, wenn  $RI-M$  raumbezogenes Wissen derart modifiziert, daß es nicht mit dem kompatibel ist, bezüglich welchem  $NLP-M$  Beschreibungen generiert. Es wird davon ausgegangen, daß zu-

	VOA-M	PS-M	RI-M	S-M	NLP-M	NV-M
VOA-M	-	s	s	n	n	n
PS-M	s	-	s	n	n	n
RI-M	s	s	-	n	s	n
S-M	n	n	n	-	n	n
NLP-M	n	s	n	n	-	n
NV-M	n	n	n	n	n	-

Tabelle 4.1: Sequentialisierungsabhängigkeiten zwischen Modulen (s: sequentiell, n: nebenläufig)

erst eine räumliche Repräsentation bestimmt werden muß, bevor eine Inhaltsstruktur konstruiert werden kann.<sup>24</sup>

Sequentialisierungsrelationen repräsentieren keinen Kontrollfluß, sondern bestimmen nur die Aktivationszusammenhänge auf lokaler Ebene zwischen jeweils zwei Modulen. Es ist Aufgabe des Steuerungsmoduls unter Beachtung der Sequentialisierungsabhängigkeiten, Module in geeigneter Weise zu aktivieren und zu de-aktivieren. Eine solche Aktivierungsreihenfolge wird durch *Phasen* festgelegt, die im folgenden beschrieben werden.

#### 4.2.1.3 Aktivierung und Deaktivierung von Phasen

Eine Phase determiniert die Reihenfolge, in der Module aktiviert werden. Mit jeder Aktivierung ist ein Zeitintervall assoziiert, welches dem Modul angibt, welcher Zeitrahmen für eine bestimmte Aufgabe zur Verfügung steht. Im vorgeschlagenen Modell werden vier Phasen der Verarbeitung unterschieden:

1.  $P_{schedule}$ : Erstellung eines Ablaufplans
2.  $P_{path}$ : Suche eines geeigneten Pfades (Etablierung von Routenwissen)
3.  $P_{object}$ : Auswahl von Landmarken nach visuellen Salienz Kriterien (Etablierung von Landmarkenwissen)
4.  $P_{nlp}$ : Generierung und Präsentation inkrementeller Wegbeschreibungen

Über den *Typ* einer Phase wird einem Modul mitgeteilt, in welchem Modus es aktiviert worden ist. Wird das Raumintegrationsmodul über die Phase  $P_{path}$  aktiviert, werden im RI-M Funktionen anders evaluiert, als wenn die Aktivierung durch  $P_{object}$  erfolgt. D.h., daß durch den Typ einer Phase Kontrollinformation an die Module weitergegeben wird, die die Evaluation von Funktionen innerhalb des Moduls beeinflusst. Über die Evaluation der Funktionen wacht alleinig das jeweilige Modul. Eine derartige Modularisierung der Verarbeitung verringert den Kontrollaufwand. Die vier Phasen, die in diesem Modell integriert sind, lassen sich in zwei Klassen aufteilen. Die erste Klasse enthält die Phase  $P_{schedule}$ , die ausschließlich kontrollflußtechnische Auswirkungen besitzt. Die zweite Klasse betrifft Phasen, die inhaltsbezogen definiert sind.

Die Phase  $P_{schedule}$  ist alleinig auf das Steuerungsmodul K-M bezogen.  $P_{schedule}$  veranlaßt das Steuerungsmodul einen neuen, an  $T_{guess}$  angepaßten, Ablaufplan zu erstellen (vgl. Phase 1).

<sup>24</sup>Ein anderer Ansatz wäre, beide Module nebenläufig zu verwenden, wodurch erreicht wird, daß eine Änderung in der Konfiguration augenblicklich die Auswahl einer anderen Menge von Inhaltsstrukturen nach sich zieht. Da aber in den Untersuchungen ein solcher Fall nicht vorkam, wird von einer sequentiellen Abfolge ausgegangen (s(RI-M, NLP-M)).

$P_{\text{schedule}}$ :	Activate-Control-Module( $T_{\text{proc}}$ );
-------------------------	---

Phase 1: Phase der Ablaufplanbestimmung und -ausführung

Die anderen Phasen beziehen sich auf inhaltliche Aspekte. In der Phase zur Auswahl von Pfadsegmenten  $P_{\text{path}}$  (vgl. Phase 2) wird zuerst die heuristische, inkrementelle Pfadsuche aktiviert, welcher  $T_{\text{ps}}$  als Zeiträumen zur Verfügung hat (Schritt 1 und 2 in Abbildung 4.11). Danach wird das  $PS-M$  deaktiviert und das  $RI-M$  aktiviert (Schritt 3). Die relevanten räumlichen Relationen werden anschließend ausgewertet (Schritte 4 und 5). Das Ergebnis wird im Zeiträumen  $T_{\text{ri}}$  durch das Raumintegrationsmodul in eine sogenannte *räumliche Konfigurationsbeschreibung* integriert (vgl. Phase 2).

$P_{\text{path}}$ :	Activate-Path-Selection-Module( $T_{\text{ps}}$ ); Activate-Spatial-Module( $T_{\text{ri}}$ );
---------------------	---

Phase 2: Phase der Pfadsuche und -integration

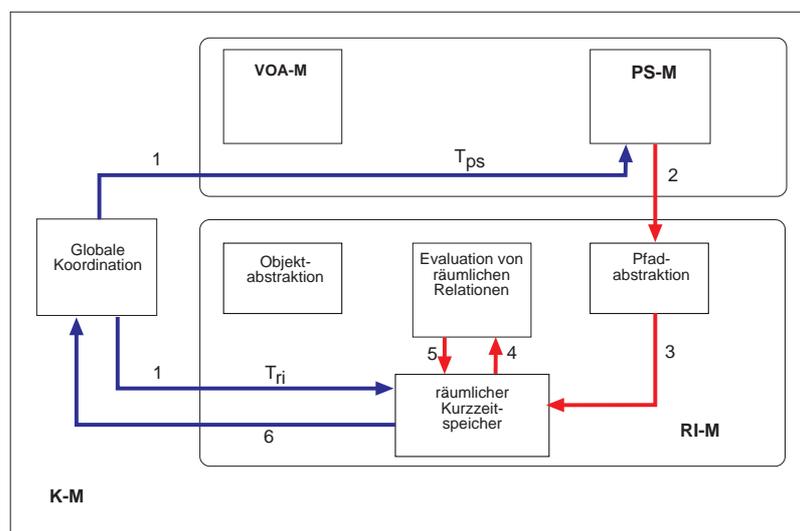


Abbildung 4.11: Evaluationschritte der Phase  $P_{\text{path}}$

Wie die empirischen Untersuchungen zeigen, verzichteten die Testpersonen in restriktiven Situationen auf die Beschreibung von Landmarken und bezogen sich ausschließlich auf Aktionen und Straßenelemente. Ist der Zeiträumen  $T_{\text{proc}}$  hinreichend groß, so wird durch die Phase  $P_{\text{object}}$  das Modul zur visuellen Objektauswahl evaluiert (vgl. Phase 3), welche referenzbildend eine

Menge von salienten Objekten selektiert (Schritte 1, 2 und 3 in Abbildung 4.12). Die Anzahl der selektierten Objekte hängt sowohl vom vorgegebenen Zeitrahmen  $T_{object}$  (vgl. Phase 3), als auch von der Situation selbst ab. Hiernach werden die Objekte, auf die gleiche Weise wie vorher die Pfadabschnitte, in die räumliche Repräsentationsstruktur durch Evaluation geeigneter räumlicher Relationen integriert (Schritte 4 und 5).

$P_{object}$ :	Activate-Visual-Module( $T_{object}$ ); Activate-Spatial-Module( $T_{ri}$ );
----------------	---

Phase 3: Phase der visuellen Objektauswahl und -integration

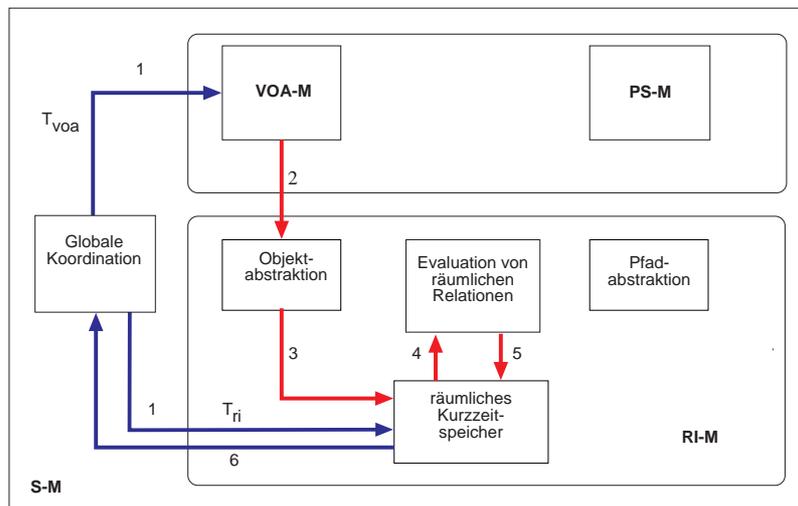


Abbildung 4.12: Evaluationsschritte der Phase  $P_{object}$

Das sprachliche Verhalten wird letztendlich über die Sprachgenerierungs- und die Präsentationsphase  $P_{nlp}$  gesteuert (vgl. Phase 4). Zuerst werden unter Berücksichtigung von  $T_{nlg}$  an die Situation und den Zielen des Agenten angepasste Inhaltsstrukturen bestimmt (Schritt 1 in Abbildung 4.13), durch Sprachfunktionen in Oberflächenstrukturen (Schritt 2) und nach der Generierung in Verbalstrukturen transformiert (Schritt 4). Anschließend wird das Präsentationsmodul aktiviert, welches darüber wacht, daß die Beschreibung im richtigen Zeitrahmen  $T_{nlp}$  gegeben wird (Schritt 5).

Im folgenden werden Ablaufpläne beschrieben, die einzelne Phasen zu komplexen Verarbeitungsschritten integrieren.

#### 4.2.1.4 Aktivierung und Evaluation von Ablaufplänen

Einzelne Phasen erlauben noch kein koordiniertes Gesamtverhalten. In einem weiteren Hierarchisierungsschritt werden die Phasen zu *Ablaufplänen* integriert. Es wird zwischen *minimalen* und *erweiterten Ablaufplänen* unterschieden. Sie definieren eine Unterscheidung in reine Aktionsbeschreibungen und solche, die durch Referenzierung von Lokationen und Objekten die

$P_{nlp}$ :      Activate-NLG-Module( $T_{nlp}$ );  
                  Activate-Presentation-Module( $T_{nlp}$ );

Phase 4: Phase der Sprachgenerierung und Präsentation

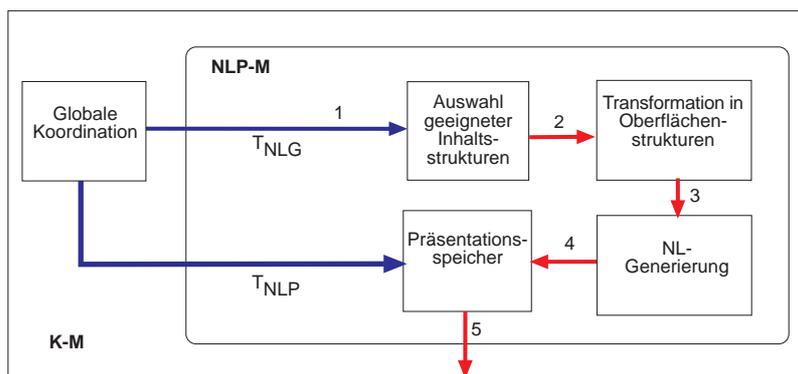


Abbildung 4.13: Evaluationsschritte bei der Ausführung von  $P_{nlp}$

Beschreibung erweitern.

Bei einer listenartigen Darstellung besteht ein Ablaufplan aus einer Liste von Teillisten (siehe Beispiel 1). Die Evaluation erfolgt sequentiell über alle Teillisten und parallel innerhalb einer Teilliste. Elemente einer Teilliste repräsentieren Phasen. Sobald eine Phase aktiviert worden ist, blockiert sie die Evaluation der nachfolgenden Teilliste des Ablaufplans. Den Phasen assoziiert sind Zeitrahmen, die angeben, wieviel Zeit der Phase zur Verfügung steht. Die nächste Teilliste wird erst dann aktiviert, wenn für alle aktivierten Phasen der Teilliste die Blockierung aufgehoben worden ist.

AP-1      (( $P_1, t_1$ ) ( $P_2, t_2$ ))  
                  ( $P_3, t_3$ )  
                  ( $P_4, t_4$ ))

Beispiel 1: Ablaufplan  $AP-1$

In Beispiel 1 werden in einem ersten Schritt die Phasen  $P_1$  und  $P_2$  aktiviert. Ein Zeitrahmen wird von einer Phase zerlegt und an die zu aktivierenden Module propagiert. Sobald die Blockierungen von  $P_1$  und  $P_2$  aufgehoben sind, werden die Phasen der nächsten Teilliste aktiviert, was im obigen Beispiel  $P_3$  ist. Es werden in einem Ablaufplan dann zwei Phasen aktiviert, wenn die zweite Phase vom Typ  $P_{schedule}$  ist, wodurch ein alternativer Ablaufplan zum aktuellen bestimmt wird. Die Aktivierung von  $P_{schedule}$  bewirkt keine Blockierung nachfolgender Phasen. Der aktuell aktive Ablaufplan ist durch die globale Variable  $*schedule*$  zugreifbar.

Ein *minimaler Ablaufplan* gewährleistet, daß der Agent sich an neue Situationen grundlegend

anpaßt. Daraus resultierende Beschreibungen enthalten die minimale Information, welche sicherstellt, daß der Rezipient eine, wenn auch minimale, so doch adäquate Beschreibung erhält. Bei einem minimalen Ablaufplan (vgl. Beispiel 2) werden zuerst geeignete Pfadabschnitte gesucht und zueinander in Beziehung gesetzt ( $P_{path}, t_{path}$ ). Nebenläufig dazu wird ein erweiterter Ablaufplan ( $P_{schedule}, t_{schedule}$ ) generiert, wie in Abbildung 4.14 illustriert ist.

$AP_{minimal}$	$((P_{path}, t_{path}), (P_{schedule}, t_{schedule}))$ $((P_{nlp}, t_{nlp}))$
----------------	--

### Beispiel 2: Minimaler Ablaufplan

Diese Verzweigungsmöglichkeit des Planungsansatzes beruht auf der Unabhängigkeit der einzelnen Module. Ist ein erweiterter Ablaufplan bestimmt worden, so wird dieser in der gleichen Weise wie der minimale abgearbeitet. Für den Fall, daß eine bereits aktive Phase erneut aktiviert werden soll, muß damit solange gewartet werden, bis diese Phase durch den vorhergehenden Ablaufplan deaktiviert wurde. Dadurch werden Überschneidungen und Inkonsistenzen vermieden.

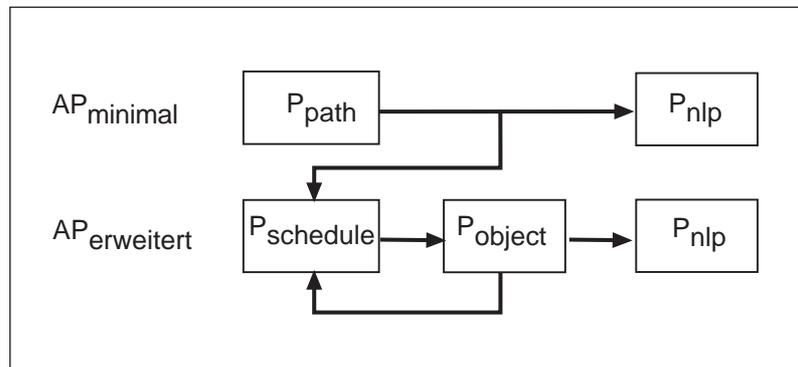


Abbildung 4.14: Nebenläufige Ausführung von minimalen und erweiterten Ablaufplänen

Ein mögliche Erweiterung des Ablaufplans besteht in der Integration der Objektauswahl-Phase ( $P_{object}, t_{object}$ ). Da während  $AP_{minimal}$  die aktuellen Pfadabschnitte bereits ausgewählt worden sind, ist es nicht notwendig  $P_{path}$  erneut zu aktivieren (vgl. Beispiel 3).

$AP_{extended}$	$((P_{object}, t_{object}), (P_{schedule}, t_{schedule}))$ $((P_{nlp}, t_{nlp}))$
-----------------	--

### Beispiel 3: Erweiterter Ablaufplan

Die Integration einer Landmarke in die räumliche Konfigurationsbeschreibung ( $P_{object}, t_{object}$ ) hat zur Folge, daß eine um eine Ortsreferenz erweiterte Inhaltsstruktur selegiert und in eine natürlichsprachliche Beschreibung umgesetzt wird. Wiederum wird nebenläufig zu  $P_{object}$  erneut  $P_{schedule}$  aktiviert, um zu überprüfen, ob weitere Landmarken integriert und damit die Beschreibung verbessert werden kann (vgl. Abbildung 4.14). Dies hängt einerseits davon ab, ob es in der Situation weitere saliente Landmarken gibt, und davon, ob noch ein genügend großer Zeitrahmen  $T_{guess}$  zur Verfügung steht.

Abschließend werden die Funktionen des Steuerungsmoduls vorgestellt, die die Steuerung des gesamten Modells innehaben.

### 4.2.1.5 Funktionen des Steuerungsmoduls

Das Verhalten des Steuerungsmoduls wird durch eine globale Schleife determiniert, die initial einen minimalen Ablaufplan bestimmt und überprüft, ob noch auszuführende Ablaufpläne vorhanden sind. Ist dies nicht der Fall, so befindet sich der Agent in einem Endzustand. Die Ausführung von Ablaufplänen erfolgt durch die Funktion *Evaluate-Schedule* (vgl. Funktion 1). Diese Funktion startet die Evaluation zur Bestimmung zeitlicher Beschränkungen (*Evaluate-Time-Frame*), die in der Funktion *Determine-Schedule* dazu verwendet werden, einen adäquaten Ablaufplan zu bestimmen (siehe Funktion 2). Die Ausführung eines Ablaufplans erfolgt durch die Funktion *Execute-Schedule* (siehe Abbildung 4.15).

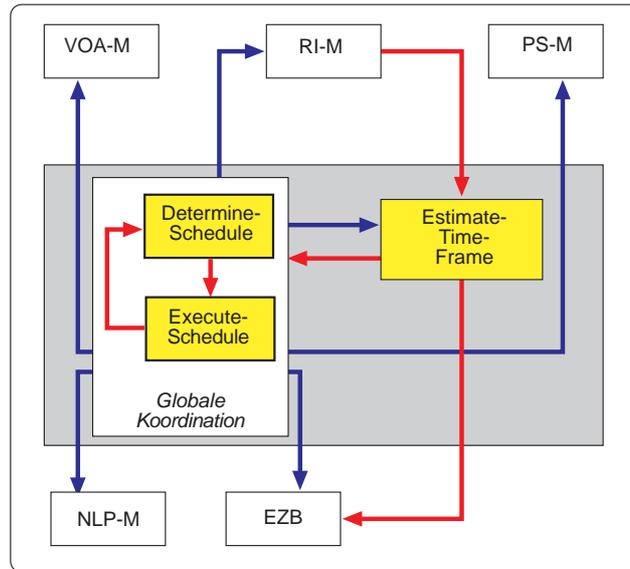


Abbildung 4.15: Aktivationszyklus zwischen den Funktionen der globalen Koordination und der Evaluation zeitlicher Beschränkungen im Steuerungsmodul

Initialisiert wird die globale Funktion *General-Control-Loop* durch einen minimalen Ablaufplan, der durch die Funktion *Determine-Minimal-Schedule* generiert wird (vgl. Funktion 2). Dieser Plan ist so konzipiert, daß der Agent ausschließlich eine minimale, auf Routenwissen basierende Beschreibung generiert. Ist der zur Verfügung stehende Zeitrahmen  $T_{guess}$  kleiner als der vom minimalen Ablaufplan geforderte, so kann keine Beschreibung generiert werden. Der Zeitrahmen  $T_{guess}$  wird durch die Funktion *Evaluate-Time-Frame* bestimmt. Dazu bestimmt *Estimate-Time-Frame* die Strecke zwischen der aktuellen Lokation des Agenten ( $loc(ego)$ )<sup>25</sup> und der Lokation des nächsten Verzweigungspunktes ( $loc(ep)$ ). Mit Hilfe der aktuellen Geschwindigkeit wird unter Verwendung der Gleichung  $t = s/v$  geschätzt, wie lange der Agent benötigt, den Verzweigungspunkt zu erreichen. Entsprechend  $T_{guess}$  wird über die Funktion *Determine-Schedule* ein geeigneter Ablaufplan ausgewählt (vgl. Funktion 1).

Entsprechend diesem Ablaufplan werden Phasen aktiviert und deaktiviert. Die Kontrollschleife wird verlassen, wenn *\*schedule\** die Nachricht *finished* enthält. Ist es nicht möglich, zu  $T_{guess}$  einen alternativen Plan zu erzeugen, geht die Funktion *Execute-Schedule* vorübergehend in einen Wartezustand, bis die Beschreibung präsentiert worden ist, die sich noch im Präsentationsspeicher befindet. Nach Ausführung der Aktion befindet sich der Agent in einer neuen Situation, was zur Folge hat, daß der minimale Ablaufplan erneut aktiviert wird (vgl. Funktion 1).

<sup>25</sup>Dies wird der Einfachheit wegen durch *ego* abgekürzt.

---

 Funktion 1: Evaluate-Schedule
 

---

```

Evaluate-Schedule(schedule)
  if (situation is not new) then
    if (for all s in schedule: for all p in phases(s): (In-Time(p))) then
      Execute-Schedule(s);
    else
      schedule := Determine-Schedule(Estimate-Time-Frame(loc(ego), loc(ep(s))));
      Execute-Schedule(schedule);
  else
    Initialize-Schedule;
    schedule := Determine-Minimal-Schedule(s);
    Execute-Schedule(schedule);
  
```

---

In konstanten zeitlichen Abständen wird überprüft, ob die einzelnen Phasen die ihnen zugeprochenen Zeitrahmen überschritten haben. Bei einer Überschreitung wird der zugehörige Ablaufplan aus der Liste *\*schedule\** entfernt. Falls der Agent sich jedoch in einer neuen Situation befindet, wird *schedule* initialisiert und ein *minimaler Ablaufplan* aktiviert (vgl. Funktion 1).

---

 Funktion 2: General-Control-Loop
 

---

```

General-Control-Loop
  schedule := Determine-Minimal-Schedule(Estimate-Time-Slot(s));
  while (finished is not in *schedule*)
  do
    Evaluate-Schedule(*schedule*);
  done
  
```

---

Sollte eine der Phasen den ihr zugeteilten Zeitrahmen überschritten haben, wird ein alternativer Ablaufplan erstellt. Im einfachsten Falle heißt dies, daß  $T_{guess}$  zu gering ist, um einen alternativen Ablaufplan zu bestimmen. In diesem Falle kann der Agent nur die Beschreibung präsentieren, die vom letzten vollständig durchlaufenen Ablaufplan generiert worden ist. Ist  $T_{guess}$  ausreichend groß, so wird ein alternativer Ablaufplan erstellt. Die aktuellen Ablaufpläne werden durch die Funktion *Execute-Schedule* ausgeführt bzw. kontrolliert (vgl. Funktion 3).

Ist das erste Element eines Ablaufplans *s* abgearbeitet worden, werden die Phasen des restlichen Ablaufplans aktiviert (vgl. Funktion 3). Befindet sich die erste zu aktivierende Phase in einem Wartezustand, so verbleibt der gesamte Ablaufplan solange im Wartestatus, bis die jeweilige Blockierung aufgehoben ist. Diese Planungsvorgehensweise ließe sich in mancher Hinsicht verbessern. Beispielsweise könnte der Zeitrahmen einer Phase dynamisch modifiziert werden. Folglich müßte ein komplexer Planungsmechanismus integriert werden, welcher evaluiert, ob und wie Ergebnisse bisheriger Berechnungen weiterverwendet werden können. Kann der Teilplan nicht weiter verwendet werden, wäre ebenso eine einfachere Strategie erfolgreich, in der eine Phase sofort deaktiviert wird, sobald ein alternativer Aufruf dieser Phase geplant worden ist. Weiterhin wäre möglich, daß das Steuerungsmodul an der inhaltlichen Planung in den Modulen beteiligt wäre. Dies wiederum würde zu Interferenzen zwischen Phasen und Modulen führen, die nur durch komplexe Schlußfolgerungsmechanismen berechnet werden können. Vor diesem Hintergrund ist der verwendete Planungsansatz einfach, besitzt aber die Eigenschaft, Kontrollwissen und inhaltsbezogenes Wissen strikt zu trennen, was den Planungsaufwand minimiert und die Planungsstrategie effizient einsetzbar macht.

---

 Funktion 3: Execute-Schedule
 

---

```

Execute-Schedule(sched)
  for all s in sched;
  case:
    (Deactive-Status(first(s))):
      s := rest(s);
      Activate-Phases(first(s));
    (Waiting-Status(first(s))):
      if (not Already-Activated-Phases(first(s))) then
        Activate-Phases(first(s));
  if (sched is not nil) and (sched is not in *schedule*) then
    *schedule* := (union *schedule* sched);
  if (not Already-Activated-Phases(first(s))) then
    Activate-Phases(first(s));
  else
    (Set-Into-Waiting-Status(first(s)));
  
```

---

Im weiteren werden die inhaltsorientierten Module vorgestellt. Dabei wird zur einfacheren Beschreibung von zeitlichen Beschränkungen und der Kommunikation mit dem Steuerungsmodul abstrahiert.

#### 4.2.2 Funktionale Beschreibung des Pfadsuchmoduls

In diesem Kapitel werden Funktionen diskutiert, die es dem Agenten erlauben, aus einer Kartenrepräsentation der Umgebung solche Pfadelemente auszuwählen, die er für die Fortsetzung der Navigation in der aktuellen Situation benötigt. Hierzu wird eine inkrementelle heuristische Pfadsuche verwendet (vgl. Abbildung 4.16). Um eine effiziente Suche zu ermöglichen, werden Fokussierungstechniken verwendet.

Als Grundlage für die Beschreibung eines Weges muß der Agent wissen, welche Sequenz von Pfadabschnitten zu verfolgen ist, um die Ziellokation zu erreichen. Die inkrementelle, heuristische Pfadsuche erfolgt in drei Schritten. Zuerst wird mittels der Pfadsuchfunktion *Incremental-Heuristic-Path-Search*, der auf den aktuellen Pfadabschnitt ( $P_{\text{pred}(vp)} \rightarrow vp$ ) folgende Pfadabschnitt ( $P_{vp} \rightarrow \text{succ}(vp)$ ) determiniert.<sup>26</sup>

Als Bedingung an eine Sequenz von Pfadabschnitten gilt, daß es in dieser Sequenz zu jedem Pfadabschnitt  $P_a$  einen weiteren Pfadabschnitt  $P_b$  gibt, der direkt angrenzt. Das wird durch eine topologische Relation ( $\#meets\#(P_a P_b)$ ) repräsentiert. Dies indiziert, daß es eine Lokation  $VP$  gibt, mit  $P_x \rightarrow vp = P_a$  und  $P_{vp} \rightarrow y = P_b$  die den Verzweigungspunkt zwischen diesen beiden Pfadabschnitten definiert (siehe Abbildung ??).

Als externe Repräsentation räumlicher Information wird eine Kartenrepräsentation der Umgebung als Grundlage für die Bestimmung eines Pfades von einem Start- zu einem Zielort verwendet. Zwischen der Umgebung und der Karte besteht eine injektive Abbildung, die jedem Pfadabschnitt eine Kante und jedem Verzweigungspunkt einen Knoten zuordnet (vgl. Abbildung 4.2). Aufgrund dieser Abbildung ist es dem Agenten möglich, einen Pfad zu suchen und diesen dazu zu verwenden, in einer Umgebung zur Ziellokation zu gelangen.

Im folgenden werden die Funktionen der inkrementellen heuristischen Pfadsuche im einzelnen vorgestellt.

---

<sup>26</sup>Dabei steht  $vp$  für Verzweigungspunkt,  $\text{succ}(vp)$  ist dem  $vp$  nachfolgende und  $\text{pred}(vp)$  der vorhergehende Verzweigungspunkt.

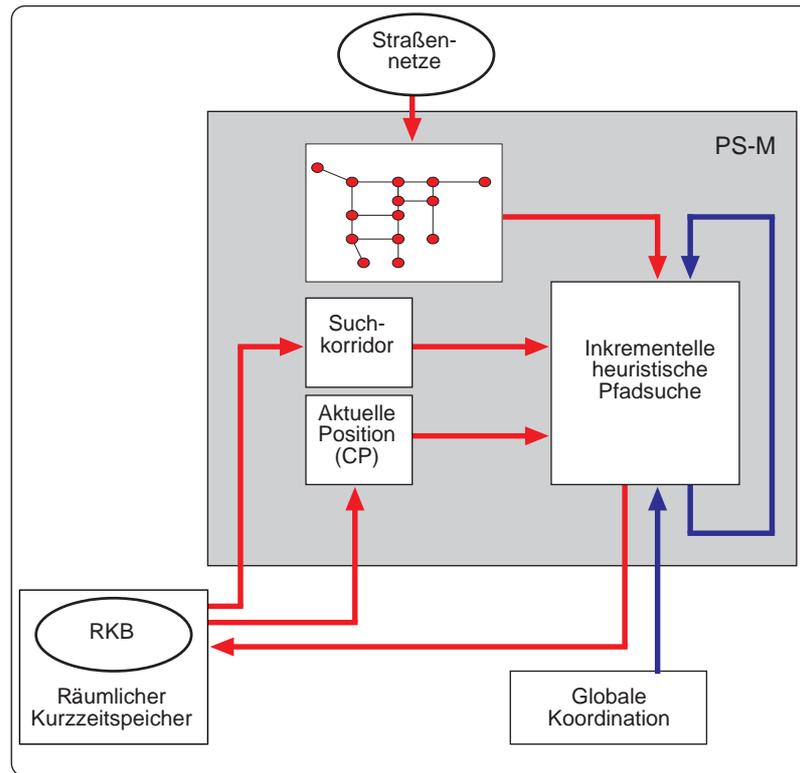


Abbildung 4.16: Inkrementelle Pfadsuche

#### 4.2.2.1 Funktionen der inkrementellen heuristischen Pfadsuche

Zur Berechnung eines geeigneten *Nachfolgepfadabschnitts* bestimmt die inkrementelle, heuristische Suchfunktion einen *kostenminimalen* Nachfolgeknoten (*move*) zum Endknoten (*state*) der aktuellen Kante ( $P_{\text{pred}(\text{state})} \rightarrow \text{state}$ ) (vgl. Abbildung 4.17). Die Kosten setzen sich dabei aus den tatsächlichen Kosten zusammen, die benötigt werden, um den Knoten *move* zu erreichen und denen, die man von *move* aus entsprechend der Schätzung zum Zielknoten benötigt. Ist ein *Nachfolgeknoten* bestimmt, so wird davon ein *Pfadabschnitt* ( $P_{\text{state}} \rightarrow \text{move}$ ) abgeleitet. Da davon ausgegangen wird, daß der Agent sich in einer für ihn neuen Umgebung befindet, wird der Pfad inkrementell mit einer begrenzten *Vorausschau* (look-ahead) bestimmt.

Für eine solche Aufgabenstellung ist die Familie der RTA\*-Algorithmen besonders geeignet. Eine Anpassung des RTA\*-Algorithmus ([Kor85b], [Kor85a], [Kor90], [RW91]) erlaubt eine inkrementelle Bestimmung von Knotensequenzen, basierend auf dem Konzept der *heuristischen Suchalgorithmen*. Die Domäne der Realumgebung gibt Beschränkungen für den abstrakten Suchraum vor. Eine Grundgröße bei der Auswahl von Pfadabschnitten ist der räumliche Abstand zwischen zwei Knoten. Die Repräsentation des abstrakten Suchraums spiegelt diese topographische Eigenschaft des Domänenraums wieder (vgl. Abbildung 4.2).

Eine Erweiterung der inkrementellen Suchalgorithmen bezüglich der Domäne der Straßenkarten, wird durch Einführung eines *räumlichen Korridors* erreicht, der eine globale Einschränkung des Suchraums angibt. Dies repräsentiert die Unterscheidung von *Routenwissen* und *Überblickswissen* (vgl. Kapitel 2.3). Durch informelle Befragungen konnte festgestellt werden, daß Probanden zwar lokal von Knoten zu Knoten, bezüglich einer domänenabhängigen Granularitätsstufe, planen, jedoch dabei auf eine grobe Richtung zum Zielort zurückgreifen. Mittels dieser Richtung sind sie in der Lage, durch minimale Information über das räumliche Gebiet, welches Start- und

Zielort enthält, den Pfad weitgehend lokal zu planen (vgl. Abbildung 4.17).

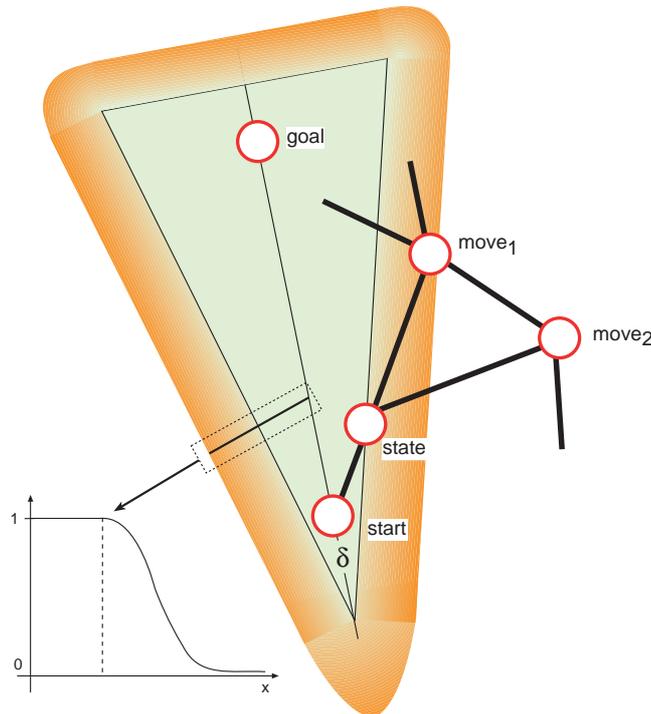


Abbildung 4.17: Korridor bezüglich des Gebiets um den Startknoten *start* und den Zielknoten *goal*

Die Funktion *Incremental-Heuristic-Path-Search* wird über  $P_{path}$  iterativ aufgerufen, solange der Zielknoten *goal* noch nicht erreicht worden ist. Über den Parameter *limit* wird angegeben, wieviele Schritte im voraus geplant werden, d.h. wieviel sukzessive Knoten und Kanten des Pfades zu einem Zeitpunkt bestimmt werden. Da kein Rückgriff auf raumbezogene Langzeitrepräsentationen verwendet wird, beträgt die Begrenzung *limit* ungefähr eins bis zwei, da dies einer Approximation der Pfadabschnitte entspricht, die über den visuellen Wahrnehmungsbereich erschließbar sind.

Der aktuelle Knoten *state* wird expandiert, indem alle direkten Nachfolgeknoten bis zu einer Tiefe von *limit* evaluiert werden (vgl. Funktion 5). Das Ergebnis dieser Evaluation sind die Kosten  $f'$ , die dem Nachfolgeknoten zugeordnet sind (vgl. Funktion 4). Alle berechneten Kosten werden in einer Hashtabelle *Table* gespeichert. Der Nachfolgeknoten von *state* ist ein solcher, welchem der geringste Kostenwert  $f'$  zugeordnet ist.

Die zentrale Funktion der Pfadsuche heißt *Evaluation*, in welcher die Kosten eines direkten Nachfolgeknotens berechnet werden (vgl. Funktion 5). In dieser Funktion wird eine Schleife solange durchlaufen, wie die Liste der aktiven Knoten (*Open*) nicht leer ist. Der minimale Kostenwert des Knotens *move* wird in der Variablen  $\alpha$  gespeichert. Entsprechend der im A\*-Algorithmus verwendeten Funktion  $f$ , setzen sich die Kosten aus den *tatsächlichen* Kosten für die Strecke von einem Knoten *state* zum Nachfolgeknoten *move* und den *heuristisch* geschätzten Kosten für das Erreichen des Zielknotens von *move* zusammen. Ausgehend von *move* werden dessen Nachfolgeknoten daraufhin überprüft, welchem bis zu einer Tiefe *limit* oder dem Erreichen des Zielknotens die geringsten Übergangskosten zugeordnet sind. Knoten, die den aktuellen Wert von  $\alpha$  unterschreiten, werden weiterverfolgt. Als Anpassung an die Domäne physikalischer Umgebungen sind die Übergangskosten (*move-cost*) zwischen einem Knoten *node* und einem Nachfolgeknoten

## Funktion 4: Incremental-Heuristic-Path-Search

```

Incremental-Heuristic-Path-Search(state, goal)
  while (state  $\neq$  goal) do
    expand state;
    for each child move of state do
      f'(move) := evaluate(move, limit);
    done;
    Table(state) = update(state, children);
    state = move with lowest f'
  done

```

*child* von der Lage im Raum und von den Eigenschaften des Pfadelementes abhängig. Unter den Eigenschaften eines Pfadelementes ist zu verstehen, ob ein Weg im aktuellen Bewegungsmodus begehbar ist, wodurch beispielsweise für einen Fahrradfahrer i.a. Autobahnverbindungen entfallen. Daraus folgt für die Fahrt mit einem Auto die Berücksichtigung der Höchstgeschwindigkeit oder aber die Häufigkeit von Staus und Autounfällen.<sup>27</sup> Als Rückgabewert liefert *Evaluation* die Kosten für den günstigsten Pfad mit maximaler Tiefe *limit* in Richtung des Zielknotens, der im Knoten *move* startet.

## Funktion 5: Evaluation

```

Evaluation(move, limit)
  Open := {move};
   $\alpha := \infty$ ;
  f(move) := g(move) + h(move);
  while Open not nil do
    node = pop Open;
    expand node;
    for all child of node do
      g(child) = g(node) + move-cost(node, child);
      f(child) = g(child) + h(child, goal);
      if f(child) <  $\alpha$  then
        if depth = limit or goal(child) then
           $\alpha = f(child)$ ;
        else
          push child on Open;
    done
  done
  return( $\alpha$ )

```

In die heuristische Berechnung der Kosten von einem Knoten zum Zielknoten wird das Konzept des *räumlichen Korridors* integriert (vgl. Funktion 6). Die Funktion *distance* gibt als Unterschätzung der Distanz des minimalen Pfades von einem Knoten *node* zum Zielknoten *goal* die Länge der direkten Verbindungslinie an. Über die Domäne des physikalischen Raumes ist jedem Knoten eine Lokation zugeordnet. Nicht jede dieser Lokationen ist als Verbindungsknoten in gleicher Weise geeignet. Knoten, die innerhalb des räumlichen Korridors liegen, werden

<sup>27</sup>Hierdurch eröffnet sich ein breites Feld für die Diskussion der Verbindung eines Suchalgorithmus und seiner Anwendungsdomäne. Dies soll hier nicht weiter diskutiert werden, da dies zum eigentlichen Problem der Wegbeschreibung nicht sonderlich viel beiträgt. Angemerkt sei aber, daß diese Domänenabhängigkeit in der Implementation dieses Modells für verschiedene Bewegungsmodi integriert worden sind (vgl. [AK96]).

bevorzugt vor solchen gewählt, die außerhalb liegen (vgl. Abbildung 4.17). Diese Abhängigkeit von der räumlichen Lage wird durch die Gewichtungsfunktion *corridor* erreicht. Das intendierte Verhalten ist durch eine monoton fallende, stetige und normalisierte Funktion modelliert. Knoten innerhalb des Korridors erhalten den Wert 1, solche außerhalb erhalten Werte zwischen 0 bis 1 (vgl. Abbildung 4.17). Die Größe des Korridors ist domänenabhängig. Wird ein Pfad bezüglich einer Fortbewegung per Automobil gesucht, so ist ein größerer Korridor zu verwenden, als wenn die Fortbewegung zu Fuß erfolgt.<sup>28</sup>

---

Funktion 6: Heuristische Funktion *h*

```

h(node, goal)
  dist = distance(node, goal) / corridor(node);
  return(dist);

```

---

Die Speicherung von Kostenwerten der Knoten dient einerseits der Effizienzsteigerung und andererseits zur Auflösung von Zyklen. Hierzu werden die Kosten, die einem Knoten zugeordnet sind, bei jedem nochmaligen Erreichen dieses Knotens um einen Wert  $\epsilon$  erhöht. Um Seiteneffekte zu vermeiden, ist der Wert von  $\epsilon$  größer Null, aber kleiner gleich den minimalen Kosten zwischen beliebigen zwei Knoten des Suchraums.

---

Funktion 7: Update

```

Update(state, children)
  if Table(state) = nil then
    return(min(f'(children)) +  $\epsilon$ )
  else
    return(Table(state) +  $\epsilon$ )

```

---

#### 4.2.2.2 Pfadabschnitte und Verzweigungspunkte

Gleichzeitig mit der Auswahl des kostengünstigsten Nachfolgeknotens *move* wird die Kante  $P_{state} \rightarrow move$  festgelegt, über die man diesen Knoten vom Endknoten *state* der aktuell aktiven Kante aus erreichen kann. Diese Kante wird auf einen in der Situation identifizierbaren Pfadabschnitt  $P_{state} \rightarrow move$  abgebildet. Ist ein Pfadabschnitt  $P_{start} \rightarrow end$  identifiziert worden, so wird er auf verschiedene Weise repräsentiert (*Abstract-Path-Segment*). Wie in späteren Verarbeitungsschritten deutlich wird, kann hierdurch eine Verarbeitung von pfadbezogenem Wissen erreicht werden. Ein Pfadabschnitt ist zum einen durch seinen Schwerpunkt (*cog*(P): center of gravity) und andererseits durch seinen Anfangs- (*start*) und Endpunkt (*end*) repräsentiert.<sup>29</sup> Ein Verzweigungspunkt *VP* wird zusätzlich durch alle Knoten, die mit den Anfangsknoten von Wegen übereinstimmen und seinem Schwerpunkt repräsentiert. Im letzten Verarbeitungsschritt der Phase  $P_{path}$  werden Pfadabschnitte in *räumliche Konfigurationsbeschreibungen* integriert (vgl. Abbildung 4.16). Hierzu werden *deiktische* und *intrinsische* räumliche Relationen zwischen dem neuen Pfadabschnitt und dem Agenten zu bereits integrierten Pfadsegmenten und zu selektierten Landmarken, die *räumlich nahe* sind, etabliert (vgl. Kapitel 4.2.4).

<sup>28</sup>Ein ellipsenförmiger Korridor ist als Modellierung besser geeignet, jedoch aufwendiger zu berechnen. Aus Effizienzgründen wird in der vorliegenden Implementation eine Approximation durch ein Dreieck verwendet. Über den Öffnungswinkel  $\delta$  wird dabei die Fortbewegungsart berücksichtigt.

<sup>29</sup>Durch die zu *start* und *end* assoziierten Lokationen ist die Richtung des Pfadabschnitts  $P_A \rightarrow B$  vorgegeben.

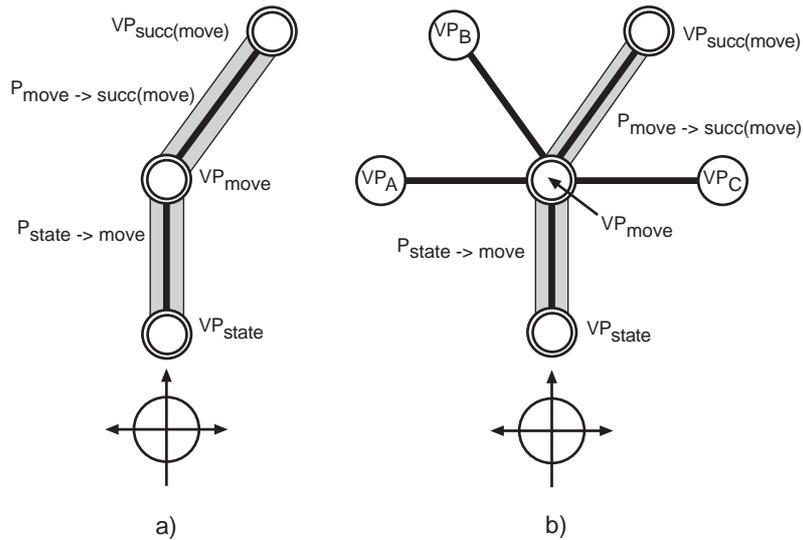


Abbildung 4.18: Ableitung eines punktförmigen Verzweigungspunkts a) ohne und b) mit alternativen Pfadabschnitten

Im weiteren wird vorgestellt, Verzweigungspunkte, und dabei insbesondere freie Plätze und Kreisverkehre, behandelt werden.

#### 4.2.2.3 Expansion von Verzweigungspunkten

Aus der Definition eines Pfadabschnitts (vgl. Definition 4) geht hervor, daß ein Übergang von einem Pfadabschnitt zum nächsten dann stattfindet, wenn sich die Ausrichtung des egozentrischen Referenzsystems qualitativ ändert. Eine Änderung dieses Referenzsystems bedingt in der Domäne der Wegbeschreibungen eine Rotation, insbesondere bezüglich der Frontalachse, in der horizontalen Ebene. Befindet sich der Agent an einer Lokation, an der eine solche Rotation stattfindet, können alternative Pfadabschnitten vorhanden sein. Auf einem *Verzweigungspunkt* verläßt der Agent die eindimensionale Struktur des Pfadabschnitts und betritt einen zweidimensionalen Raum, in dem er sich in zwei Dimension bewegen kann. Ein Verzweigungspunkt selbst, soweit es kein Platz oder Kreisverkehr ist, wird als punktförmig angenommen. Die Bestimmung der Knoten *move* und *succ(move)* etabliert gleichzeitig Verzweigungspunkte  $VP_{move}$  und  $VP_{succ(move)}$  (siehe Abbildung 4.18). Für die Pfadabschnitte  $P_{state} \rightarrow move$  und  $P_{move} \rightarrow succ(move)$  müssen die topologischen Relationen gelten, damit sie begehbar sind ( $\#meets\# P_{state} \rightarrow move P_{move} \rightarrow succ(move)$ ), mit  $VP_{move} \in P_{state} \rightarrow move \cap P_{move} \rightarrow succ(move)$ .

Es gibt Verbindungen zwischen zwei Pfadabschnitten, die nicht einfach durch eine punktförmige Abstraktion repräsentiert werden können. In städtischen Umgebungen sind dies *freie Plätze*, aber auch Artefakte, wie beispielsweise Kreisverkehre. Auf der Ebene des Suchraums werden freie Plätze durch *freie Verzweigungspunkte* repräsentiert. Genau wie andere Verzweigungspunkte, indizieren sie Verbindungen zwischen Pfadabschnitten. Ein freier Verzweigungspunkt gibt keine Pfadabschnitte vor. Auf der Straßenkarte ist ein freier Platz als punktförmig angegeben, so daß die Pfadsuche als Ergebnis die Sequenz  $(P_{state} \rightarrow move P_{move} \rightarrow succ(move))$  liefert (vgl. Abbildung 4.19a).

Für eine Beschreibung dieser Pfadabschnitte reicht diese Repräsentation nicht aus. Es wird im weiteren davon ausgegangen, daß zwischen den Pfadabschnitten  $P_{state} \rightarrow move$  und  $P_{move} \rightarrow succ(move)$  ein freier Platz ( $VP_{move}$ ) liegt, bezüglich dem auf Grund sei-

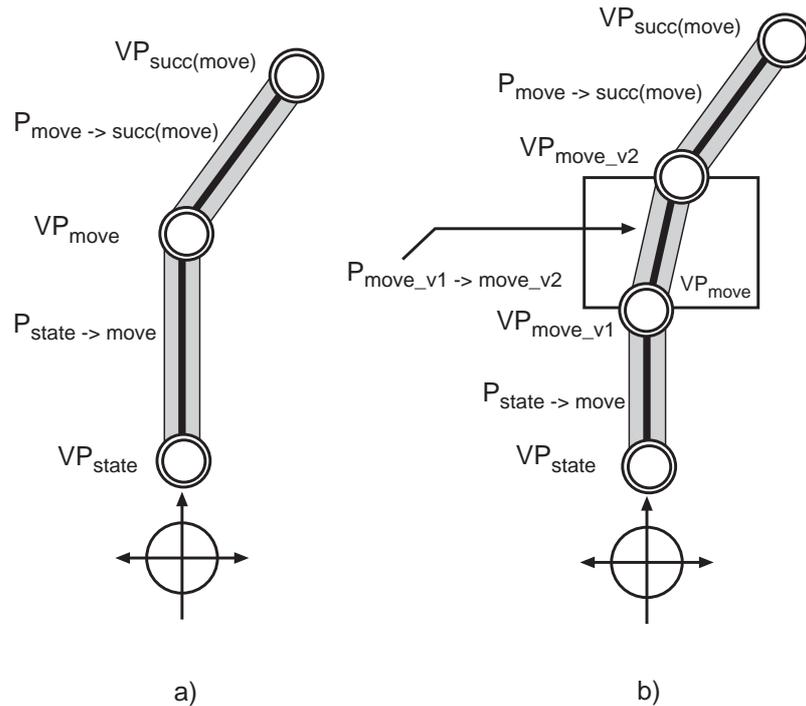


Abbildung 4.19: Expansion eines freien Platzes. Eine punktförmige Repräsentation aus der Kartenrepräsentation (a)) wird in eine expandierte überführt (b))

ner physikalischen Ausdehnung ein eigener Pfadabschnitt spezifiziert werden muß, der es dem Hörer erlaubt, den anschließenden Pfadabschnitt  $P_{move} \rightarrow succ(move)$  zu erreichen. Dazu wird die Fläche, die durch einen freien Platz gegeben wird, in weitere Verzweigungspunkte und Pfadabschnitte dekomponiert. Der Eintrittsort in den freien Platz  $VP_{move}$  ist gleichzeitig der Endpunkt des Pfadabschnitts  $P_{state} \rightarrow move$ . Der Knoten  $VP_{move}$  wird durch einen Expansionsschritt in eine Knoten-Kante-Knoten-Sequenz transformiert ( $P_{state} \rightarrow move$   $P_{move} \rightarrow succ(move)$ ) in ( $P_{state} \rightarrow move$   $P_{move_{v1}} \rightarrow move_{v2}$   $P_{move} \rightarrow succ(move)$ ). Die beiden Verzweigungspunkte  $VP_{move_{v1}}$  und  $VP_{move_{v2}}$  werden durch einen *virtuellen* Pfadabschnitt  $P_{move_{v1}} \rightarrow move_{v2}$  verbunden, der keine direkte Entsprechung in der Knoten- bzw. Kantenrepräsentation der Karte besitzt (vgl. Abbildung 4.19b).

Eine Strategie, die die direkte Verbindungslinie verwendet, ist nicht in jedem Falle möglich. Bedingt durch Konventionen und anderen Abhängigkeiten kann die Bestimmung eines virtuellen Pfades bezüglich eines Platzes verschiedene Strategien erfordern. Dazu soll der Fall eines Kreisverkehrs als Beispiel dienen. In den meisten Ländern ist es Konvention, den Kreisverkehr rechtsherum zu durchfahren, in manchen Ländern jedoch linksherum. Diese Konvention oktroyiert eine ausgezeichnete Strategie auf die räumliche Struktur des Kreisverkehrs.

$$\begin{array}{l}
 (P_{state} \rightarrow move \ P_{move} \rightarrow succ(move)) \\
 \Rightarrow \\
 (P_{state} \rightarrow vp_{v1} \ P_{vp_{v1}} \rightarrow vp_{v2} \ P_{vp_{v2}} \rightarrow vp_{v3} \\
 P_{vp_{v3}} \rightarrow vp_{v4} \ P_{vp_{v4}} \rightarrow VP_{succ(move)})
 \end{array}$$

Beispiel 4: Beispiel des expandierten Ablaufplans eines Kreisverkehrs

Ergibt die Padsuche folgende Abfolge der Pfadelemente ( $P_{state} \rightarrow move$   $P_{move} \rightarrow succ(move)$ )



Neben Information über Pfadabschnitte werden in inkrementellen Wegbeschreibungen visuell perzipierte Objekte referenziert. Geeignete Funktionen zur Auswahl von Objekten werden im folgenden Kapitel diskutiert.

### 4.2.3 Funktionen des Objektauswahlmoduls

Die Verankerung von Wegbeschreibungen in einer Situation erfolgt vor allem durch Referenzierung auf visuell-perzipierte Objekte bzw. Lokationen, die einen gewissen Grad an Auffälligkeit besitzen. Auffälligkeit besitzt ein Objekt durch unterschiedliche visuelle Merkmale, von denen *Farbe*, *Höhe* und *Breite* zu den wichtigsten zählen (vgl. [Man86] und Kapitel 2). Unter der Annahme, daß sich Auffälligkeiten von Merkmalen typspezifisch bestimmen lassen, wird im weiteren ein formales Modell zur merkmalsbasierten Berechnung von *visuellen Salienzen*, als Formalisierung der visuellen Auffälligkeit eines Objektes, vorgestellt.

Zwei Ebenen sind bei der Analyse von visuellen Salienzen zu unterscheiden. Auf einer *Objektrepräsentationsebene* werden solche Objekte verwaltet, die in einer gegebenen Situation von einem bestimmten Blickwinkel aus sichtbar sind (vgl. rechter Teil in Abbildung 4.21). Auf einer zweiten Ebene wird eine *Rasterrepräsentation* durch Überlagerung eines Rasters auf die Projektionsebene generiert (vgl. linker Teil in Abbildung 4.21). Jeder einzelnen Lokation auf dem Raster sind *rekursive, visuelle Merkmale* zugeordnet, die an das jeweils zugehörige Objekt vererbt werden (vgl. Kapitel 2). Diese Vererbung ist solange einfacher Natur, wie die Verteilung visueller Merkmale über assoziierte Objekte homogen ist.<sup>30</sup> Der Fokus dieses Kapitels liegt auf der Diskussion eines Verfahrens, welches unter der Homogenitätsannahme bezüglich *rekursiver* und *linearer* visueller Merkmale visuelle Salienzen von Objekten berechnet. Das dadurch modellierte Objektauswahlmodul stellt eine abstrakte, referenzkonstituierende Schnittstelle zwischen visueller Wahrnehmung und raumbezogenem Wissen her, worüber raumsprachliche Beschreibungen zu Entitäten einer Situation Bezug nehmen.<sup>31</sup> Insgesamt geht es vorrangig um eine Bewertung der visuellen Auffälligkeit von Objekten als Grundlage zur Referenzbildung sprachlicher Ausdrücke und weniger um eine Modellierung visueller Wahrnehmungsfunktionen.

#### 4.2.3.1 Allgemeine Beschreibung der Auswahl visuell salienter Objekte

Die globale Funktion, welche visuell saliente Objekte in einer Situation aus einem bestimmten Blickwinkel in einer Situation  $S$  bestimmt, ist *Select-Salient-Object* (vgl. Funktion 8). Visuelle Salienzen dienen als Filter, um unter einer Vielzahl von Objekten solche auszuwählen, die auffällig sind. Dies bedingt einerseits eine minimale Analyse der gesamten visuell-zugänglichen Information einer Situation und andererseits eine explizite Analyse geeigneter Objekte. Dieser "trade-off" wird durch zeitliche Beschränkungen beeinflusst.

Der erste Schritt besteht in der Bestimmung von Objekten auf einer zweidimensionalen Repräsentationsebene durch die Funktion *Determine-Visible-Objects*. Sie stellt die Schnittstelle zur visuellen Objekterkennung her (vgl. Abbildung 4.22). Als Approximation der sichtbaren Flächen eines Objektes wird ein minimal-umschreibendes Rechteck auf der Projektionsfläche verwendet. Zur Analyse der Situation wird die visuelle Information auf eine zweidimensionale Ebene parallel projiziert, welche sich vertikal zur Bewegungsrichtung befindet (*Project-Situation*) (vgl. Abbildung 4.23). Durch die Projektionsebene wird eine Repräsentation der Situation generiert, die

---

<sup>30</sup>Ist die Homogenitätsannahme nicht gewährleistet, ist eine komplexere Formalisierung zu berücksichtigen, die dem Objekt integrierte oder partiell definierte rekursive Merkmale zuordnet, wie es beispielsweise bei einem regenbogenfarbigen Objekt der Fall ist.

<sup>31</sup>Ob die Information über die visuellen Merkmale eines Objektes durch visuelle Wahrnehmungsprozesse oder durch einen Zugriff auf ein digitalisiertes Objektmodell erhalten wird, ist in diesem Ansatz von sekundärem Interesse.

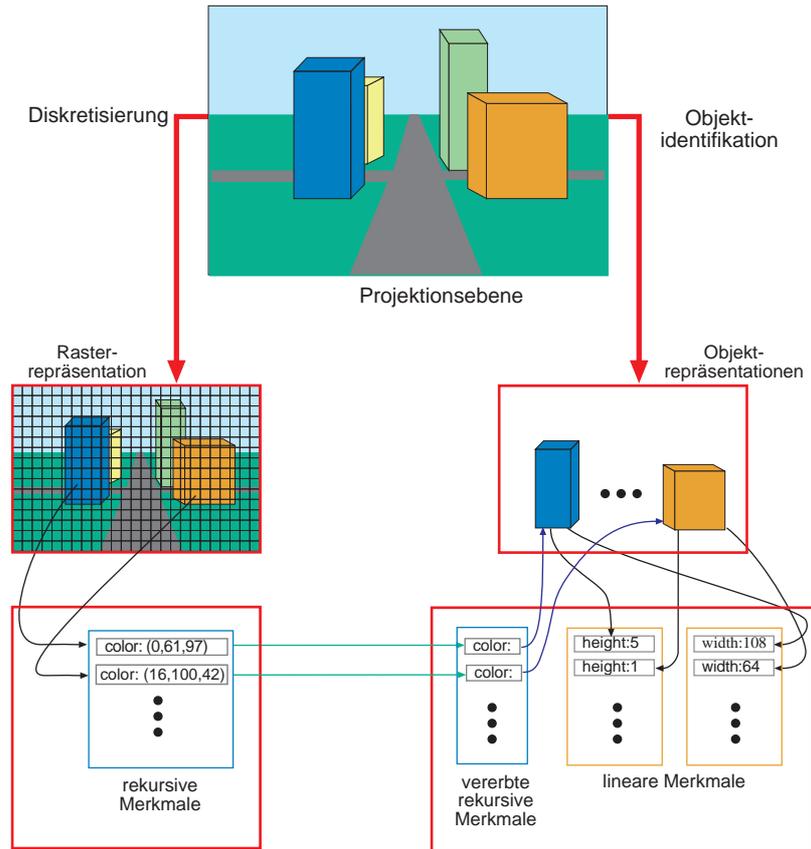


Abbildung 4.21: Zerlegung einer perspektivischen Sicht in lokationsbezogene und objektbezogene Repräsentationen

vergleichbar ist mit der retinalen Projektion. Auf der Projektionsebene wird, entsprechend den psychologischen Befunden zur Aufmerksamkeit, bezüglich einer Lokation (z.B. [Bro58], [KT84], [PIFC87]) oder mehrerer Lokationen (z.B. [Ull85], [Pyl89], [CMD<sup>+</sup>91], [HR94]) ein räumlicher Fokusbereich bestimmt. Der *räumliche Fokusbereich* determiniert den Bereich des Sichtfeldes, auf den ein Agent höchste Aufmerksamkeit richtet, wobei dies wesentlich von den aktuellen Intentionen des Betrachters abhängig ist. In Bezug auf inkrementelle Wegbeschreibungen ist die Ausrichtung des räumlichen Fokus durch *pfadbezogene Intentionen* bestimmt (vgl. Kapitel 4.2.4.3). Aus diesem Grunde ist die Funktion *Determine-Fokus* sowohl durch die Situation als auch durch die Intentionen des Agenten parametrisiert (vgl. Funktion 9). Nachfolgend wird die Repräsentation mit Hilfe eines *Rasters* in *Zellen* zerlegt. Dadurch wird die Objektrepräsentationsebene verlassen und für *rekursive* Merkmale eine Rasterrepräsentation aufgebaut (vgl. Abbildung 4.21). Die Größe einer Zelle gibt an, wie genau die Analyse einer Situation erfolgt. Aus neuropsychologischen Experimenten aber auch formalen Überlegungen (vgl. [OK95], [OAV93], [Tso90]) folgt, daß Probanden sich entsprechend der gegebenen Situation und ihrer Intentionen über die Genauigkeit der Auflösung an die visuellen Daten anpassen. Jede Zelle des Feldes erbt die rekursiven Merkmale des ihr zugeordneten Objektes (*raster-rep* in Funktion 8 und Abbildung 4.21). Im darauffolgenden Verarbeitungsschritt wird, entsprechend der “feature integration theory” (vgl. Kapitel 2), für die betrachteten Merkmalstypen (Höhe, Breite und Farbe) eine eigene Rasterrepräsentation angelegt, welche als *Merkmalskarte* bezeichnet werden. Diese Dekomposition ist unter der Annahme möglich, daß Merkmale unabhängig voneinander verarbeitet und später wieder integriert werden können (vgl. Kapitel 2).

## Funktion 8: Select-Salient-Object

```

Select-Salient-Object(S, intention, Tobject)
  Determine-Visible-Objects;
  sit := Project-Situation(Situation);
  Determine-Focus(S, intention, Tobject);
  raster-rep := Discretize-Into-Raster(sit);
  f-maps := Decompose-Into-Feature-Maps(raster-rep);
  s-maps := Determine-Saliency-In-F-Maps(f-maps);
  i-map := Integrate-Feature-Maps(s-maps);
  salient-locs := Determine-Salient-Locs(i-map);
  salient-objects := Determine-Salient-Objects(salient-locs);

```

Das Ergebnis der Funktion *Decompose-Into-Feature-Maps* ist die Dekomposition in typspezifische Merkmalskarten (f-maps). Auf jeder dieser Merkmalskarten werden saliente Lokationen bestimmt, die angeben, ob ein bestimmtes Merkmal im Verhältnis zur gesamten Situation salient ist (*Determine-Saliency-In-F-Maps*). Unterschieden wird die Bestimmung von visuellen Salienzen rekursiver und linearer Merkmale sowie die lokationserhaltende Integration dieser in der *globalen Merkmalskarte*. Die merkmaltypspezifische Berechnung visueller Salienzen rekursiver Merkmale ist in Kapitel 4.2.3.6 für das Merkmal „Farbe“ und linearer Merkmale in Kapitel 4.2.3.7 für die Merkmale „Höhe“ und „Breite“ explizit angegeben. Auf Grund der Rekursionseigenschaft werden visuelle Salienzen rekursiver Merkmale auf einer Rasterebene und solche für lineare Merkmale auf Objektebene berechnet. Das Resultat ist die Zuordnung eines Salienzwertes einer Zelle relativ zu einer Merkmalskarte. Überlagert man die einzelnen typspezifischen Merkmalskarten und integriert die Salienzwerte einer jeden Zelle in topographieerhaltender Weise (*Integrate-Feature-Maps*), ergibt sich eine merkmalsübergreifende Repräsentation (i-map) salienter Zellen (*Determine-Salient-Locs*). Dieser Verarbeitungsschritt befindet sich immer noch auf der Ebene der Rasterrepräsentation. Erst im darauffolgenden Schritt wird die Verbindung zur Objektebene hergestellt, wenn die Referenz von Objekten zu Salienzen einzelner Zellen etabliert wird (*Determine-Salient-Objects*). Das Ergebnis der Funktion *Select-Salient-Landmarks* sind solche Objekte einer Situation, deren Salienzwert über einem ressourcenabhängigen Schwellwert liegen. Insgesamt ergibt sich ein funktionaler Zusammenhang, wie er in Abbildung 4.22 illustriert ist. Im weiteren sollen die Funktionen im einzelnen genauer diskutiert werden.

#### 4.2.3.2 Identifikation sichtbarer Objekte

Die allgemeine Bestimmung von visuell-sichtbaren Objekten ist eine komplexe Problemstellung, die hier nur soweit diskutiert werden soll, wie sie für diese Arbeit relevant ist. Das Ergebnis einer vollständigen Identifikation von Objekten ist der Aufbau von multiplen, unterschiedlich abstrahierten Objektrepräsentationen (vgl. [Her96]). Auf der untersten Stufe der Komplexität ist die Schwerpunktrepräsentation, die alleinig aus einem Punkt bestimmt ist. Daneben sind u.a. zweidimensionale Repräsentationen in der horizontalen Ebene und minimal-umschreibende Quader definiert. Auf einer weitaus komplexeren Ebene befinden sich Repräsentationen von Haupt- und Nebenachsen des Objektes sowie vollständige geometrische Modellbeschreibungen (z.B. [Mar82], [Bie90]).

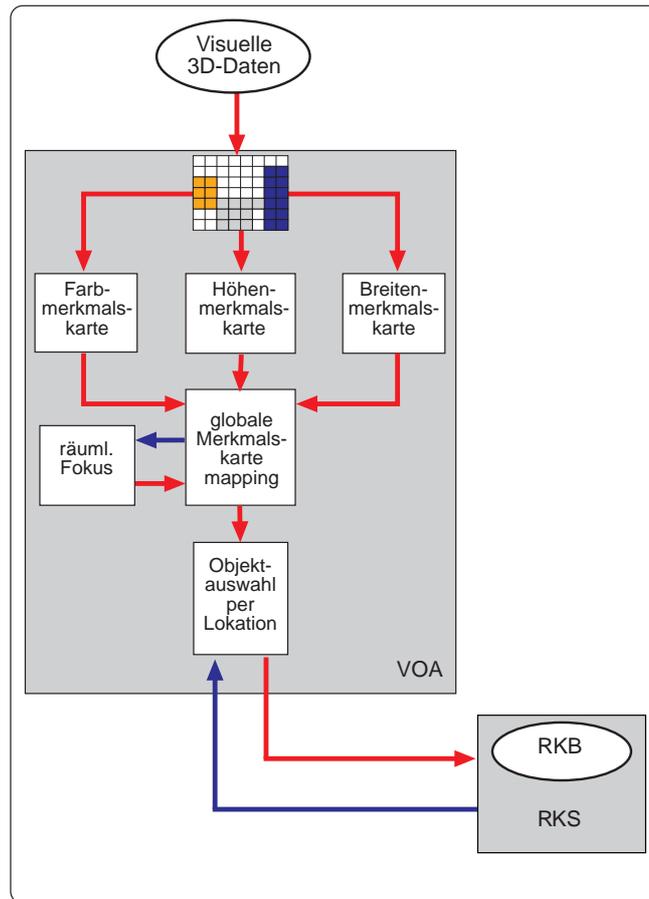


Abbildung 4.22: Der funktionale Aufbau des visuellen Objektauswahlmoduls

#### 4.2.3.3 Abbildung der perspektivischen Sicht auf die Projektionsebene

Visuelle Salienzen sind in dem hier vorliegenden Modell bezüglich eines gegebenen Zeitpunkts unter einem bestimmten Blickwinkel des Agenten definiert. Die Dreidimensionalität der Umgebung, in der sich der Agent bewegt, spielt bei dieser Berechnung der visuellen Salienz eine untergeordnete Rolle. Da Tiefeninformation nur einen geringen Beitrag bei der Anordnung von Objekten in einer Situation besitzt, wird die aktuelle Sicht auf eine transversal zur Blickrichtung definierte Projektionsebene abgebildet (vgl. Abbildung 4.23).

Objekte, die auf die Projektionsebene abgebildet werden, sind mit den Flächen ihrer Projektion assoziiert. Die gesamte sichtbare Fläche eines Objektes wird durch ein minimal-umschreibendes Rechteck approximiert. Diese Approximation ist solange hinreichend, wie nur solche Objekte betrachtet werden, deren Konvexität oder Konkavität in einem zu vernachlässigenden Rahmen bleiben. Das Ergebnis der Funktion von *Project-Situation* ist eine zweidimensionale Projektion der Situation.

#### 4.2.3.4 Berücksichtigung der räumlichen Fokussierung

Bezüglich dieser Projektionsebene fokussiert der Agent auf Teilbereiche einer Situation. Eng mit dem Konzept eines Fokus ist das Gebiet der *selektiven Aufmerksamkeit* verbunden. Wie zahlreiche empirische Erhebungen zeigen, ermöglicht eine Beschränkung der Aufmerksamkeit, daß der Zusammenhang zwischen kognitivem Verarbeitungsaufwand und Komplexität der Situation

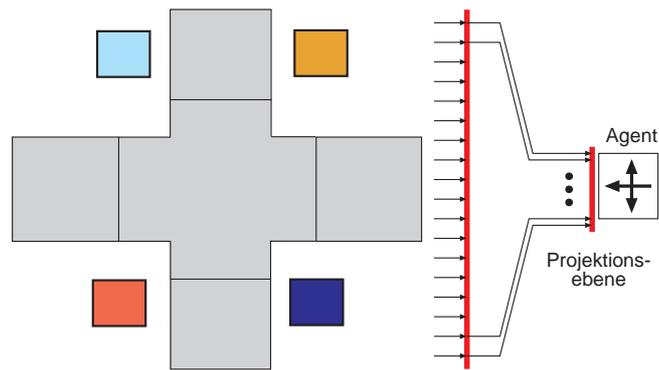


Abbildung 4.23: Parallelprojektion der Situation auf die Projektionsebene

nicht linear ist. Vielmehr erlauben selektive Aufmerksamkeitprozesse, daß selbst in komplexen Situationen noch adäquate Entscheidungen getroffen werden können.<sup>32</sup>

#### Funktion 9: Determine-Focus

*Determine-Focus*( $S$ , *intention*,  $T_{object}$ )  
 Determine-Center-Of-Focus(*intention*);  
 Determine-Area-Of-Focus( $T_{object}$ );

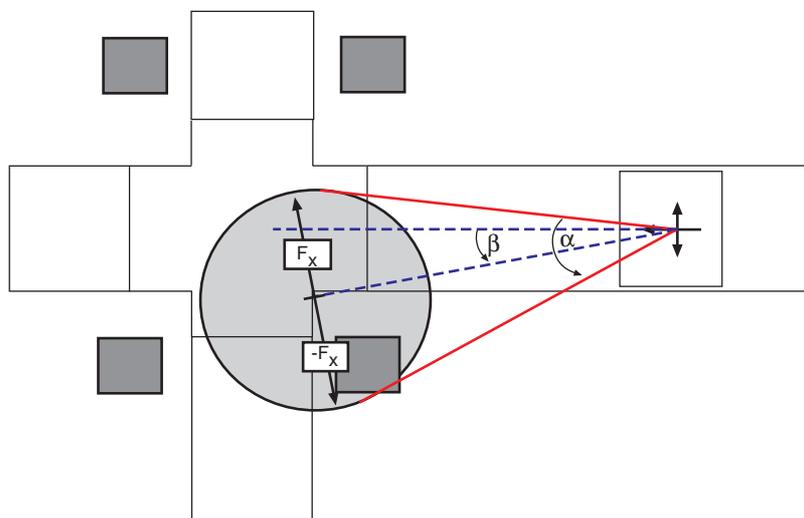


Abbildung 4.24: Ausrichtung des räumlichen Fokus entsprechend pfadbezogener Intention, nach links abbiegen zu wollen, und Berücksichtigung des Öffnungswinkels  $\alpha$  und Abweichungswinkels  $\beta$ .

Wie aus den empirischen Untersuchungen hervorgeht (vgl. Kapitel 3), werden vor allem solche Bereiche fokussiert, die in der Nähe des nächsten Verzweigungspunktes und dabei tendenziell in Richtung der intendierten Aktion liegen. Entsprechend dem aktuell verwendeten Referenzsystem (vgl. Kapitel 4.2.4.1) ändert sich der Abweichungswinkel  $\beta$  von der Hauptbewegungsrichtung bei

<sup>32</sup>Der räumliche Fokus findet in der Pfadsuche im Konzept des *Korridors* eine Entsprechung (vgl. Kapitel 4.2.2).

einem intendierten Abbiegevorgang (vgl. Abbildung 4.24). Objekte außerhalb des Fokusbereichs bleiben nahezu unbeachtet. Der Öffnungswinkel  $\alpha$  ist eine Funktion des Zeitrahmens  $T_{object}$ . Je kleiner  $T_{object}$  ist, desto kleiner ist der Öffnungswinkel  $\alpha$ . Unter der Annahme eines rechteckigen Fokusbereichs (vgl. Abbildung 4.25) ergeben sich direkte Abhängigkeiten zwischen  $T_{object}$  und der x- und y-Komponente des Fokusbereichs  $F$  ( $F_x, F_y$ ) (vgl. Gleichung 1).

$$(1) \quad F_x = K_x * T_{object}; F_y = K_y * T_{object}$$

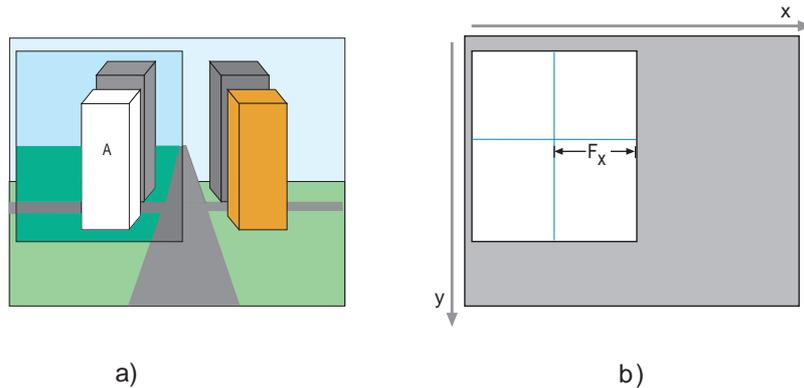


Abbildung 4.25: a) Perspektivische Sicht, b) Variabler räumlicher Fokusbereich mit Breite  $2 * F_x$

Unter Annahme eines linearen Zusammenhangs sind die Koeffizienten  $K_x$  und  $K_y$  an Hand von informellen Befragungen bestimmt worden.  $F_x$  und  $F_y$  determinieren den Bereich, in dem alle Objekte höchste Aufmerksamkeit besitzen.

#### 4.2.3.5 Abbildung der Projektions- auf die Lokationsebene

Die Eigenschaft rekursiver Merkmale bedingt, daß deren Träger beliebig unterteilt werden können, ohne daß einer der Teile den Wert dieses Merkmal verliert. Die Berechnung visueller Salienzen auf der Basis rekursiver Merkmale verlangt dadurch eine hinreichend feine Zerlegung des Trägers, die es effizient ermöglicht, Anteile rekursiver Merkmale zu berechnen. Hierzu geeignet ist die objektunabhängige Rasterzerlegung. Dazu wird die kontinuierliche Repräsentation einer perspektivischen Sicht durch Überlagerung eines Raster diskretisiert (vgl. den Diskretisierungsschritt in Abbildung 4.26). Sichtbare Objektflächen stimmen i.a. nicht mit den Grenzen überein, die durch das Raster vorgegeben sind (vgl. mittlere Graphik in Abbildung 4.26). Jede sichtbare Fläche wird soweit expandiert, bis sie mit einem minimal umschreibenden Rastergebiet übereinstimmt (vgl. Approximationsschritt in Abbildung 4.26). Je feiner die Rasterung, d.h. je kleiner eine einzelne Zelle ist, um so genauer ist die Approximation. Auch dies führt zu einem "trade-off" zwischen der Komplexität der Berechnung und der Qualität der Approximation<sup>33</sup>. Die Werte der rekursiven Merkmale, die den Zellen zugeordnet werden, entsprechen den Mittelwerten der ihnen zugeordneten Gebiete auf der Projektionsebene.<sup>34</sup>

Der Fehler, der durch die Approximation einer Fläche  $A$  bedingt ist, setzt sich bezüglich der

<sup>33</sup>Informell ist eine solche Vorgehensweise von Toda in seinem "Fungus Eater" vorgeschlagen worden (vgl. [Tod62]).

<sup>34</sup>Dies korrespondiert mit dem Phänomen, daß eine Menge dünner, schwarzer, eng nebeneinanderliegender Linien einen grauen Farbton erzeugen. Bei homogener Merkmalsverteilung eines Objekts ist diese Zuordnung trivial. Sie wird jedoch relevant, wenn Objekte heterogene rekursive Merkmale besitzen.

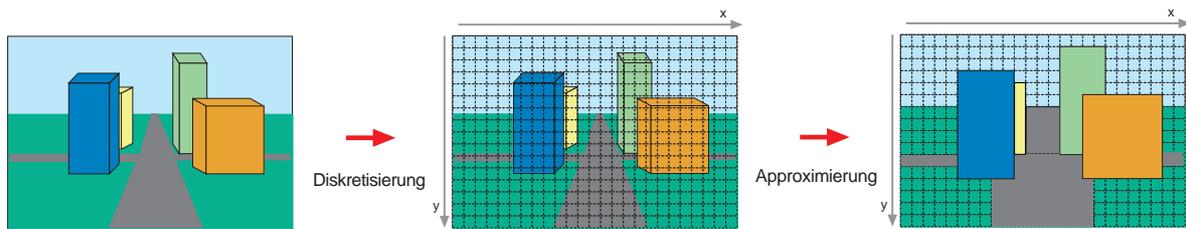


Abbildung 4.26: Transformation einer Situationsprojektion in eine diskrete Lokationsrepräsentation

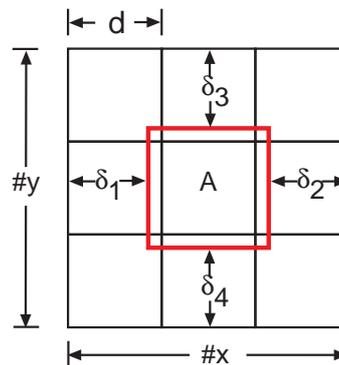


Abbildung 4.27: Größenfehler, die durch die Approximation der Fläche durch das minimal umschließende Rastergebiet entstehen.

x-Achse  $f_x$  und y-Achse  $f_y$  aus den Distanzen zur nächsten Zelle zusammen (vgl. Gleichungen 2 und Abbildung 4.27).

$$(2) \quad \begin{aligned} f_x &= \delta_1 + \delta_2 \\ f_y &= \delta_3 + \delta_4 \end{aligned}$$

Der Fehler ist im schlechtesten Falle in jeder Dimension kleiner als die Ausdehnung einer einzelnen Zelle  $d$  (vgl. Ungleichung 3).

$$(3) \quad f_x < d; f_y < d$$

Daraus folgt für den *Gesamtfehler*  $f$  einer Fläche  $A$ , daß dieser nicht größer als das Quadrat der Ausdehnung eines Rasters mal der Anzahl der umschließenden Zellen ist (vgl. Gleichung 4).

$$(4) \quad f = 2 * ((\#x * f_x) + (\#y * f_y)) < 2 * ((\#x + \#y)) * d^2$$

Daraus ergibt sich ein *relativer Fehler*  $\delta f$  (vgl. Ungleichung 5). Der relative Fehler  $\delta f$  wird entsprechend den gegebenen zeitlichen Beschränkungen gewählt. Dazu gilt, daß die Größe des

Fehlers in einem umgekehrt-proportionalem Verhältnis zum Berechnungsaufwand steht.

$$(5) \quad \delta f = \frac{f}{2^{* \# x + 2^{* \# y}} \leq d^2$$

Als Ergebnis liefert Funktion *Discretize-Into-Raster* eine approximierete Rasterrepräsentation, in der jeder Zelle ein Objekt auf der Projektionsebene zugeordnet ist (vgl. Abbildung 4.21).

#### 4.2.3.6 Bestimmung visueller Salienzen bezüglich des rekursiven Merkmals „Farbe“

In welcher Weise visuelle Salienzen rekursiver Merkmale berechnet werden können, wird am Beispiel des Merkmals *Farbe* erläutert. Das rekursive Merkmal *Farbe* ist für die Etablierung visueller Salienzen von besonderer Bedeutung, da es durch empirische Studien als dominantes Merkmal bei der Auswahl von Objekten belegt ist (vgl. [Man86]).

Es gibt verschiedene Farbwahrnehmungsmodelle, die in unterschiedlicher Weise approximieren, wie der Mensch ein Farbmerkmal wahrnimmt. Beispiele für solche Modelle sind das Munsell System (vgl. [WS82]), das “natural color system” (NCS) (vgl. [Hår66]), das DIN System (vgl. [DIN80]) und das OSA/UCS System (vgl. [Der91]). Eine typische Zerlegung einer Farbe erfolgt, entsprechend der Heringschen Farbtheorie, in Farbwert, Helligkeit und Sättigung. Problematisch ist an diesen Zerlegungen, daß es selbst für das Konzept des Farbwertes keine einheitliche, formal klar spezifizierte Definition gibt (vgl. [Sv96]). Weiterhin sind die resultierenden Farbwahrnehmungsmodelle derart komplex, daß sie für einen Realzeiteinsatz nicht zu verwenden sind. Um festzustellen in wieweit sich ein Farbmerkmal in einer Situation abhebt, muß ihm zunächst ein numerischer Farbwert zugeordnet werden. Ein Farbmerkmal läßt sich in verschiedener Weise eindeutig zerlegen.

Eine allgemein gebräuchliche Zerlegung von Farben basiert auf der Verwendung der CIE-Chromatiken-Repräsentation (vgl. [WS82]). Das CIE-Chromatiken-Diagramm fußt auf einem X-Y-Z Koordinatensystem mit dem Vorteil, daß im Gegensatz zum RGB-System keine negativen Werte auftreten. Zwischen beiden Systemen bestehen Transformationsgleichungen (vgl. [WS82, S. 269]). Unter der Annahme, daß eine konstante Sättigung und eine konstante Helligkeit vorliegt, ist es ad hoc vorstellbar, den euklidischen Abstand im CIE-Chromatiken-Diagramm zu verwenden. Ein solches Vorgehen ist dann gerechtfertigt, wenn ein Betrachter zwischen allen Farbwerten in diesem Diagramm in konstanter Weise unterscheiden könnte. Wie aber McAdam zeigen konnte, treten erheblich Unterschiede beim Vergleich von Farbwahrnehmungen auf (zitiert in [WS82]).<sup>35</sup> Bezogen auf das CIE-Chromatiken-Diagramm, bestimmte McAdam in diesen Untersuchungen elliptische Bereiche um 25 Referenzpunkte, in denen die Betrachter keinen Farbunterschied feststellen konnten (siehe Abbildung 4.28, [WS82, S. 521]).

Die Ellipsen, die McAdam bestimmt hat, lassen sich über ihren Durchmesser entlang der Haupt- und Nebenachsen in drei Gruppen einteilen (vgl. Abbildung 4.28). Approximativ repräsentieren diese Gebiete Farbzerlegungen, die einen identischen Farbeindruck erzeugen. Durch Verhältnisbildung über die Größe der Ellipsen ergibt sich, daß Differenzen im „blauen“ Gebiet am besten wahrgenommen werden. In Relation zum „blauen“ Gebiet werden im „roten“ Gebiet erst bei einer 1,7-fach und im „grünen“ Gebiet bei einer 6-fach größeren Farbdifferenz Unterschiede festgestellt. Um Farbwerte über Komponenten hinweg miteinander vergleichen zu können, dienen diese Approximationen der Normierung, die die unterschiedlichen Größenverhältnissen der

<sup>35</sup>Bezogen auf einen chromatischen Punkt untersuchte McAdam, ab wann der Betrachter einen minimal wahrnehmbaren Farbunterschied feststellen kann. Als Ergebnis fand er heraus, daß der Betrachter erst einen Unterschied feststellt, wenn der Vergleichsfarbwert einen bestimmten Abstand vom Referenzpunkt annimmt.

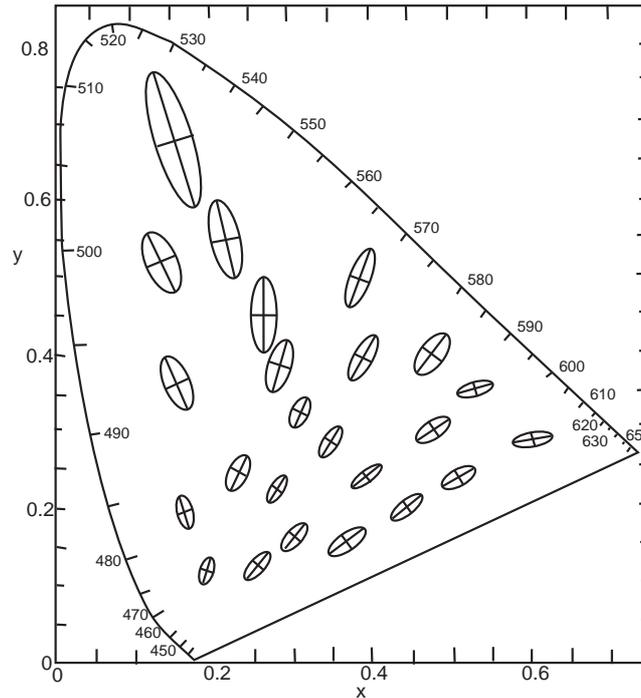


Abbildung 4.28: CIE-Diagramm mit eingezeichneten McAdam-Ellipsen (nach [WS82])

Ellipsen von McAdam repräsentieren ( $S_r = 1,7$ ;  $S_g = 6,0$ ;  $S_b = 1$ ).<sup>36</sup> Durch diese Anpassung des RGB-Modells an Farbwahrnehmungsphänomene beim Menschen, ist in approximativer Weise ein effizientes, kognitiv adäquates Berechnungsmodell für die Bestimmung der Auffälligkeit von Farbmerkmalen möglich. Der Vorteil der Verwendung des angepaßten RGB-Modells ist, daß jedes computerbasierte Farbmodell auf das RGB-Modell zurückgreift.<sup>37</sup>

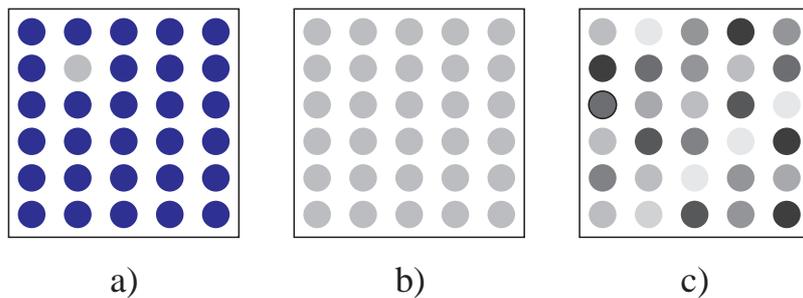


Abbildung 4.29: Veränderung von Salienzeigenschaften eines Merkmals in verschiedenen Situationen

<sup>36</sup>Bei unterschiedlichen Gesamtlichtverhältnissen unterliegt die Wahrnehmung von Farbe jedoch starken Schwankungen, sodaß Erkenntnisse, die durch Laborexperimente bei standardisierten Randbedingungen gewonnen worden sind, nicht ohne weiteres in Realumgebungen übertragen werden können. Phänomene wie der Farbkonstanzeffekt werden in diesem Ansatz nicht berücksichtigt.

<sup>37</sup>Das von Gapp implementierte Modell (vgl. [Gap95b]) beschränkt sich bei der Berechnung von visuellen Salienzen auf die Analyse von Kontrastrelationen auf Schwarz-Weiß-Graphiken. Dieses Modell kann als Erweiterung des hier vorgeschlagenen verstanden werden.

**4.2.3.6.1 Die Funktionen zur Bestimmung visueller Salienzen des Merkmals „Farbe“** Um feststellen zu können, ob ein Farbmerkmal einer Zelle als salient zu bezeichnen ist, muß definiert werden, wie sich Farbunterschiede feststellen lassen. Zur Ermittlung eines Farbunterschieds ist es notwendig, einen farbbezogenen *Referenzvektor* bezüglich einer Situation festzulegen. Ist ein Referenzvektor bestimmt, wird der Farbwert der jeweiligen Zelle dazu in Beziehung gesetzt.

Die Schwierigkeit, die sich ergibt, läßt sich am Beispiel eines roten Farbpunkts unter einer Vielzahl von blauen Farbpunkten illustrieren (vgl. Abbildung 4.29a). Der rote Farbpunkt<sup>38</sup> ist in dieser Situation salient, wohingegen er in einer Situation mit lauter anderen roten Punkten keinerlei Auffälligkeit besitzt (vgl. Abbildung 4.29b). Die Differenz eines Farbmerkmals zu einem Referenzvektor hängt von der Gesamtverteilung der Farbwerte ab. Komplizierter wird diese Frage, wenn es mehr als zwei Farbwerte in einer Situation gibt (vgl. Abbildung 4.29c). Eine solche Situation erfordert eine effiziente Bestimmung eines geeigneten Referenzvektors, der die Mischung aus allen Farbwerten liefert. Der *Referenzvektor* einer Situation,  $C_{ref}(S)$ , berechnet sich komponentenweise aus dem arithmetischen Mittel über die Farbwerte aller Zellen  $c_{ij}$  (vgl. Gleichung 6).<sup>39</sup>

$$(6) \quad C_{ref}^{\vec{}}(S) = \frac{1}{(N * M)} \begin{pmatrix} \sum_{i=1}^N \sum_{j=1}^M red(c_{ij}) \\ \sum_{i=1}^N \sum_{j=1}^M green(c_{ij}) \\ \sum_{i=1}^N \sum_{j=1}^M blue(c_{ij}) \end{pmatrix}$$

Der relative Farbanteil  $color(c_{ij})$  eines Farbwertes wird durch den Abstand zum Referenzvektor berechnet. Der resultierende Differenzvektor  $\Delta color^{\vec{}}(c_{ij})$  ist an die Verzerrungen der Farbwahrnehmung angepaßt und repräsentiert den Unterschied zwischen dem Referenzwert  $C_{ref}(S)$  und dem Farbvektor einer Zelle  $color(c_{ij})$  (vgl. Gleichung 7).

$$(7) \quad \Delta color^{\vec{}}(c_{ij}) = \|C_{ref}^{\vec{}}(S) - color^{\vec{}}(c_{ij})\| = \begin{pmatrix} \left( \frac{|red(C_{ref}(S)) - red(color(c_{ij}))|}{S_r} \right) \\ \left( \frac{|green(C_{ref}(S)) - green(color(c_{ij}))|}{S_g} \right) \\ \left( \frac{|blue(C_{ref}(S)) - blue(color(c_{ij}))|}{S_b} \right) \end{pmatrix}$$

Der Abstand zwischen zwei Farbwerten hängt von der Verteilung in der gesamten Situation ab. In einer Situation, in der eine Vielzahl von unterschiedlichen Farbwerten sichtbar sind, ist das Merkmal *Farbe* unbedeutender, als in einer solchen mit geringer Farbvariation (vgl. beispielsweise Abbildung 4.29a,c). Repräsentiert wird dieser Sachverhalt durch die Bestimmung der *lokalen Verdichtung* eines Farbmerkmals einer Zelle  $c_{ij}$  (vgl. Gleichung 8). Die Verdichtung

<sup>38</sup>In Abbildung 4.29 ist der rote Farbpunkt der Deutlichkeit halber, durch einen weißen Kreis mit schwarzen Rand dargestellt.

<sup>39</sup> $c_{ij}$  ist der Farbvektor, der einer Zelle an der Rasterposition  $(i,j)$  des Rasters zugeordnet ist;  $red(x)$ ,  $green(x)$  und  $blue(x)$  sind die Zugriffsfunktionen auf die Komponenten des Farbvektors bei RGB-Zerlegung. In dem hier vorgestellten Ansatz werden konstante Gesamtlichtverhältnisse angenommen.

ist ein Maß für die Konzentration eines Farbwertes bezüglich einer Situation.

$$(8) \quad \text{lokdens}(\vec{\text{color}}(c_{ij})) = \frac{1}{\Delta \text{red}(\text{color}(c_{ij})) + \Delta \text{green}(\text{color}(c_{ij})) + \Delta \text{blue}(\text{color}(c_{ij}))} * \begin{pmatrix} \text{red}(\Delta \vec{\text{color}}(c_{ij})) \\ \text{green}(\Delta \vec{\text{color}}(c_{ij})) \\ \text{blue}(\Delta \vec{\text{color}}(c_{ij})) \end{pmatrix}$$

Die einzelnen Farbkomponenten werden nicht unabhängig voneinander verarbeitet, sondern werden in  $\text{lokdens}(\vec{\text{color}}(c_{ij}))$  integriert. Bei monochromatischen Farbwerten besitzt genau eine Komponente den Maximalwert 1 und ist somit identisch mit einer RGB-Repräsentation. Im Falle eines gleichen Anteils aller Komponenten, was dem Farbeindruck *grau* entspricht, ist  $\text{lokdens}(\vec{\text{color}}(c_{ij}))$  gleich dem Vektor  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  und indiziert, daß keiner Farbkomponente ein besonderes Gewicht zukommt. Je stärker der Farbwert sich einer monochromatischen Farbe annähert, um so größer wird der Wert in einer Komponente.

Die Berechnung der *lokalen Verdichtungen* erfolgt für die Zellen des gesamten Sichtfeldes. Sie reflektieren den minimalen Berechnungsaufwand, der bezüglich der gesamten Farbverteilung in einer Situation erforderlich ist. Aus den lokalen Verteilungen wird mittels Integration die *globale Verdichtung* einer Situation  $\text{globdens}(\vec{\text{color}}(S))$  abgeleitet (vgl. Gleichung 9).  $N$  und  $M$  sind die Anzahl der Rasterzellen in z- und y-Ausrichtung. Je geringer die maximale lokale Verdichtung der Merkmalswerte in einer Situation ist, um so geringer ist die Wahrscheinlichkeit, daß das Merkmal *Farbe* in dieser Situation eine Relevanz für die Auffälligkeit eines Objektes beisteuert. Dies kann am Beispiel des Farbkreises verdeutlicht werden. Ein einzelner Farbwert des Farbkreises besitzt in diesem nahezu keine besondere Auffälligkeit. Im Falle eines perfekten Farbkreises ist die lokale Verdichtung über die gesamte Situation gesehen minimal, da es keine ausgezeichnete Farbkomponente gibt. Es gibt ebenso viele Anteile der „roten“, wie der „grünen“ und der „blauen“ Komponenten. Besteht die Situation jedoch nur aus „grauen“ Farbwerten, so ist die globale Verdichtung ebenfalls minimal, da keinerlei Aussage über die Verteilung der Farben gemacht werden kann.<sup>40</sup> Das andere Extrem sind monochromatische Verteilungen. In einer vollständig „rot“ gehaltenen Situation ist die globale Verdichtung in der roten Komponente maximal und in den anderen Null (vgl. Abbildung 4.29b). Zwischen diesen beiden Extrema bestimmt die globale Verdichtung, inwieweit die Farbwerte verteilt sind und in welchen Komponenten sich Maxima befinden. Dadurch, daß die globale Verdichtung auf der Integration von lokalen Verdichtungen beruht, werden nur Farbdifferenzen berücksichtigt, die auch eine wahrnehmbare Auffälligkeit in der Situation besitzen. Die Verbindung von lokaler und globaler Verdichtung ist gleichsam ein Übergang von einer atomistischen zu einer holistischen Repräsentation rekursiver Merkmale.

$$(9) \quad \text{globdens}(\vec{\text{color}}(S)) = \frac{1}{N * M} \begin{pmatrix} \sum_{i=1}^N \sum_{j=1}^M \text{lokdens}(\text{red}(c_{ij})) \\ \sum_{i=1}^N \sum_{j=1}^M \text{lokdens}(\text{green}(c_{ij})) \\ \sum_{i=1}^N \sum_{j=1}^M \text{lokdens}(\text{blue}(c_{ij})) \end{pmatrix}$$

Der Unterschied zu einer einfachen Mittelwertbildung ist, daß  $\text{globdens}(\vec{\text{color}}(S))$  die lokalen Verdichtungen und farbwahrnehmungsbedingte Verzerrungen berücksichtigt. Die globale Verdichtung repräsentiert den Farbeindruck der gesamten Situation. Durch die Bestimmung der

<sup>40</sup>Eine Weiterführung des hier vorgestellten Ansatzes ist die Bestimmung von gebietsbezogenen Verdichtungen. Eine interessante, empirische Fragestellung ist, welche „cluster“ von Testpersonen erkannt werden und wie sie den Farbkreis in Gebiete einteilen.

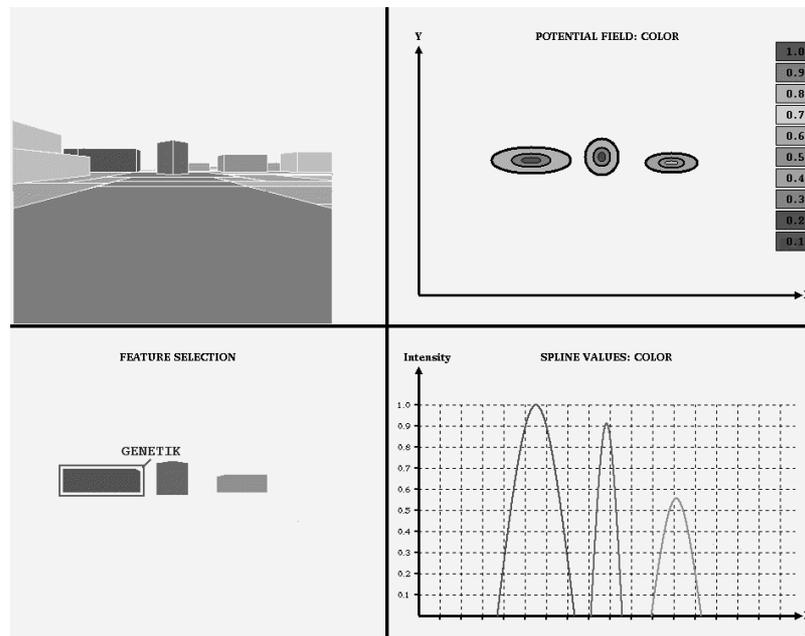
globalen Verdichtung einer Situation können die Farbmerkmale einzelner Zellen verglichen werden. Um die visuelle Salienz bezüglich des Merkmals *Farbe* einer Zelle zu bestimmen, wird der relative Farbanteil ( $\Delta\text{color}(\vec{c}_{ij})$ ) mit der globalen Verdichtung  $\text{globdens}(\vec{\text{color}}(S))$  in Beziehung gesetzt, woraus der Salienzvektor ( $\text{sal}(\vec{\text{color}}(\vec{c}_{ij}))$ ) hervorgeht (vgl. Gleichung 10). Dazu wird der relative Farbvektor der Lokation mit der *globalen Verdünnung* ( $\vec{e} - \text{density}(\vec{\text{color}}(S))$ ) multipliziert.<sup>41</sup> Anzumerken ist, daß sowohl die lokale, als auch die globale Verdichtung bezüglich normierten RGB-Werten definiert sind. Dadurch wird ein Vergleich von beiden zur Bestimmung der visuellen Salienz des Farbmerkmals einer Zelle möglich.

$$(10) \quad \text{sal}(\vec{\text{color}}(\vec{c}_{ij})) = \begin{pmatrix} \Delta(\text{red}(\Delta\vec{\text{color}}(\vec{c}_{ij})) * (1 - \text{red}(\text{globdens}(\vec{\text{color}}(S)))) \\ \Delta(\text{green}(\Delta\vec{\text{color}}(\vec{c}_{ij})) * (1 - \text{green}(\text{globdens}(\vec{\text{color}}(S)))) \\ \Delta(\text{blue}(\Delta\vec{\text{color}}(\vec{c}_{ij})) * (1 - \text{blue}(\text{globdens}(\vec{\text{color}}(S)))) \end{pmatrix}$$

Durch Betragsbildung ergibt sich der skalare, visuelle Salienzwert des Farbwertes einer Zelle ( $\text{sal}(\text{color}(\vec{c}_{ij}))$ ) (vgl. Gleichung 11).

$$(11) \quad \text{sal}(\text{color}(\vec{c}_{ij})) = |\text{sal}(\vec{\text{color}}(\vec{c}_{ij}))| / \hat{s},$$

mit  $\hat{s} = \max|\text{sal}(\text{color}(\vec{c}_{ij}))|$



[Bestimmung des salientesten Farbmerkmals][color-map-serra-sgi]

Visuelle Salienzen von Zellen werden an assoziierte Objektrepräsentationen auf der Projektionsebene zurückgeleitet (vgl. Abbildung 4.21). Dies erfolgt unter der Annahme, daß in der vorliegenden Domäne die Objekte homogene Farbverteilungen besitzen.<sup>42</sup> Somit ist der Farbwert eines Objektes gleich dem einer ihm zugeordneten Zellen  $c_{ij}$ , sodaß die Berechnung von Salienzen auf

<sup>41</sup>Die *globale Verdünnung* ist das Inverse zur *globalen Verdichtung*.

<sup>42</sup>Ist ein Objekt nicht homogen gefärbt, wird der Mittelwert über alle assoziierten Zellen bestimmt. Dieses

einer lokationsgebundenen Ebene eine Bedeutung auf Objektebene erhält. Die eher holistische Betrachtungsweise auf Rasterebene besitzt dadurch Einfluß auf die atomistische Betrachtungsweise der Objektebene.

#### 4.2.3.7 Bestimmung visueller Salienzen der linearen Merkmale „Höhe“ und „Breite“

Bei der Verarbeitung von linearen Merkmalen, wie *Höhe* und *Breite*, werden Salienzen nicht auf einer Rasterebene berechnet, da lineare Merkmale direkt an Objekte gebunden sind. Als Repräsentationsgrundlage für *Höhe* und *Breite* werden minimal umschreibende Rechtecke der Objekte auf der Projektionsebene verwendet, welche die maximale Ausdehnung eines Objektes bezüglich der horizontalen und vertikalen Dimension repräsentieren (vgl. den unteren Teil der Abbildung 4.26). Im weiteren wird, basierend auf der Invarianzannahme, davon ausgegangen, daß die Distanz und somit die Höhe und die Breite eines Objektes vom Agenten hinreichend genau abgeschätzt werden kann.

Für die Bestimmung visueller Salienzen von Höhen- und Breitenmerkmale, wird nicht auf die *absolute* Höhe und Breite eines Objektes, sondern vielmehr auf den sichtbaren Anteil auf der Projektionsebene Bezug genommen. Im Gegensatz zu rekursiven Merkmalen lassen sich *Höhe* und *Breite* durch jeweils eine einzige Dimension und demzufolge skalar repräsentieren. Der Referenzwert eines linearen Merkmals in einer Situation ergibt sich aus dem arithmetischen Mittel (vgl. die Gleichungen 12, 13).

$$(12) \quad \bar{H} = 1/n \sum_{i=1}^n \text{height}(O_i)$$

$$(13) \quad \bar{W} = 1/n \sum_{i=1}^n \text{width}(O_i)$$

Die Skalare  $\text{height}(O)$  und  $\text{width}(O)$  repräsentieren die Werte der Höhe und der Breite des korrespondierenden Rechtecks eines Objektes  $O$ . Um zu vermeiden, daß Objekte, die eine zu geringe Höhe und Breite besitzen, berücksichtigt werden, werden als Erweiterung der “feature integration theory” Schwellwerte eingeführt. Als Randbedingung für die Schwellwerte  $th_{\text{height}}$  und  $th_{\text{width}}$  gilt, daß alle Merkmalswerte größer als der typspezifische Schwellwert sein müssen (vgl. Ungleichung 14).

$$(14) \quad \text{height}(O_i) > th_{\text{height}} \wedge \text{width}(O_i) > th_{\text{width}}$$

Durch die Invarianzannahme gestützt, wird für die Bestimmung des Salienzwertes bezüglich der Höhen- und Breitenmerkmale eine lineare Metrik verwendet. Die visuellen Salienzen bezüglich des Merkmalstyps *Höhe* und *Breite* bilden sich aus der positiven Distanz zwischen dem Merkmalswert ( $\text{height}(O)$  bzw.  $\text{width}(O)$ ) und dem Referenzwert ( $\bar{H}$  bzw.  $\bar{W}$ ) (vgl. die Gleichungen

---

Verfahren ist dann nicht ausreichend, wenn das Objekt über Farbhäufungen verfügt, wie z.B. bei einem Bild von Miro. Für solche Farbverteilungen ist zu überlegen, ob die Funktion Farbsalienz  $\text{sal}(\text{color}(\text{object}))$  angewendet werden kann. Erste Versuche zeigen, daß ein solch rekursiver Aufruf zu vielversprechenden Ergebnissen führt. Eine mögliche Vorgehensweise ist es, die Farbeigenschaft des salientesten Teilbereichs des Objektes als den Farbsalienzwert des Objektes zu vererben. Liegt der maximale Farbsalienzwert unterhalb einer empirisch zu bestimmenden Schwelle, so wird die Farbsalienz des Objektes über den Farbmittelwert berechnet.

15, 16).

$$(15) \quad \text{sal}(\text{height}(O)) = |\bar{H} - \text{height}(O)|$$

$$(16) \quad \text{sal}(\text{width}(O)) = |\bar{W} - \text{width}(O)|$$

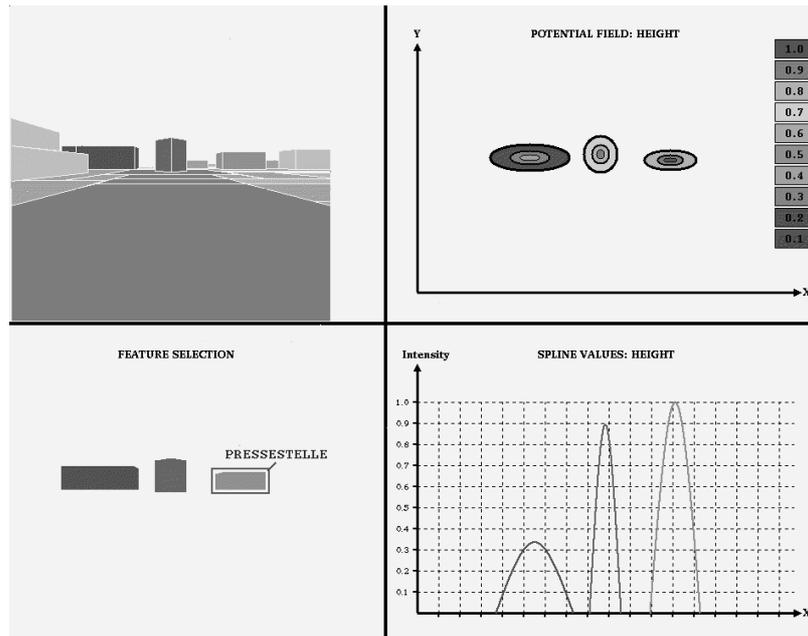


Abbildung 4.30: Bestimmung des salientesten Höhenmerkmals

In Abbildung ?? ist ein Beispiel für die Bestimmung der visuellen Salienzen bezüglich Farbmerkmalen gegeben. In diesem Falle zeigt sich, daß das Gebäude der *Genetik* diesbezüglich eine maximale visuelle Salienz besitzt. Jedoch ist der Unterschied zum nächsten Objekt nicht besonders groß. Die Berechnung der visuellen Salienzen der Breiten- und Höhenmerkmale (vgl. Abbildung 4.30, 4.31) zeigt, daß jeweils die Serra-Plastik bzw. das Gebäude der Pressestelle ein Maximum erreicht. Erst die lokationsgebundene Integration über alle Merkmalstypen hinweg ergibt, daß die Serra-Plastik die größte visuelle Salienz in der gegebenen Situation besitzt (vgl. Abbildung 4.32).

#### 4.2.3.8 Integration in globale Merkmalskarten

Das strukturkonstituierende Element bei der Berechnung visueller Salienzen sind räumliche Bereiche auf der Projektionsebene. Diesen Bereiche sind Objekte assoziiert, worüber lineare und rekursive Merkmale miteinander verbunden sind (vgl. Abbildung 4.21). Nachdem Salienzwerte eines Objektes für einzelne Merkmalstypen bestimmt worden sind, werden sie entsprechend ihrer Lokation integriert, worüber die visuelle Salienz eines Objektes determiniert ist. Die Frage, die sich stellt, ist, wie Merkmale verschiedenen Typs miteinander kombiniert werden können. Anne Treisman macht im Rahmen der “feature integration theory” darüber keine Aussagen (vgl.

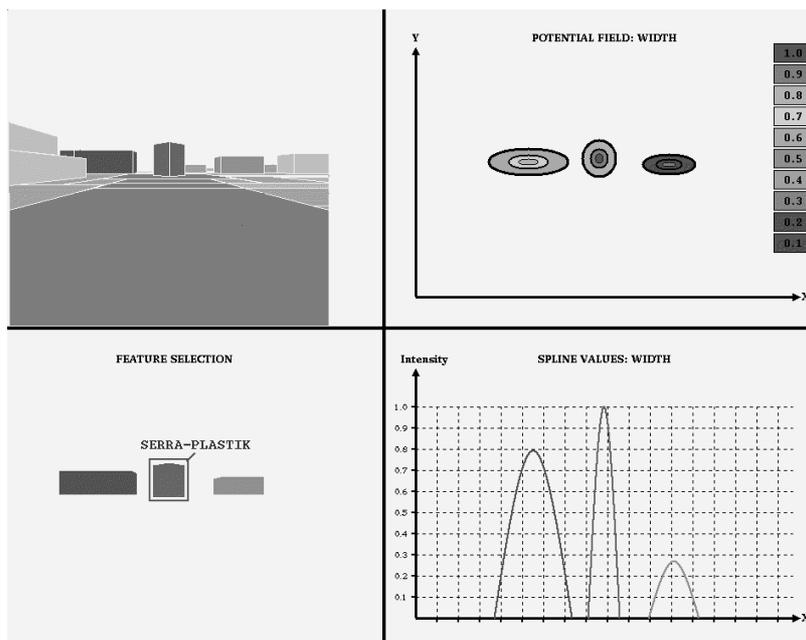


Abbildung 4.31: Bestimmung des salientesten Breitenmerkmals

Kapitel 2). Allgemein wird die Frage nach der visuellen Salienz, *vis-sal*, eines Objektes durch Gleichung 17 expliziert.

$$(17) \quad \text{vis-sal}(O) = \alpha \otimes \text{sal}(\text{color}(O)) \oplus \beta \otimes \text{sal}(\text{height}(O)) \oplus \gamma \otimes \text{sal}(\text{width}(O))$$

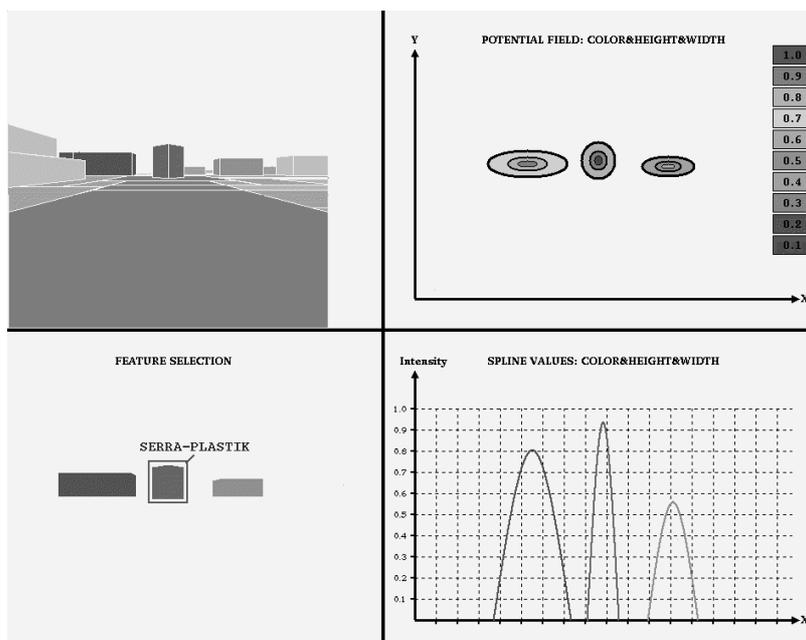


Abbildung 4.32: Bestimmung des salientesten Objektes über die globale Merkmalskarte

Sowohl die Koeffizienten der einzelnen Salienzwerte, als auch die Operation,  $\odot$  und  $\otimes$ , sind bisher nicht genauer empirisch untersucht worden. Mangels Evidenzen wird hier eine Linearkombination verwendet. Dazu sind die Koeffizienten empirisch bezüglich der simulierten 3D-Umgebung bestimmt worden (vgl. Kapitel 3). Es ergaben sich die Koeffizienten  $\alpha = 0,5$ ,  $\beta = 0,3$  und  $\gamma = 0,2$ .<sup>43</sup>

$$(18) \quad \text{vis-sal}(O) = \alpha * \text{sal}(\text{color}(O)) + \beta * \text{sal}(\text{hight}(O)) + \gamma * \text{sal}(\text{width}(O))$$

Als Ergebnis der Salienzberechnung bezüglich einer Situation werden sichtbare Objekte nach Salienzwerten geordnet und in die Liste *salient-objects* eingetragen. Nachdem den sichtbaren Objekten im Fokusbereich visuelle Salienzwerte zugeordnet worden sind, ist die aktuelle Sicht hinreichend analysiert, um eine effiziente räumliche Vorstellung der Situation aufzubauen (vgl. Abbildung 4.32).

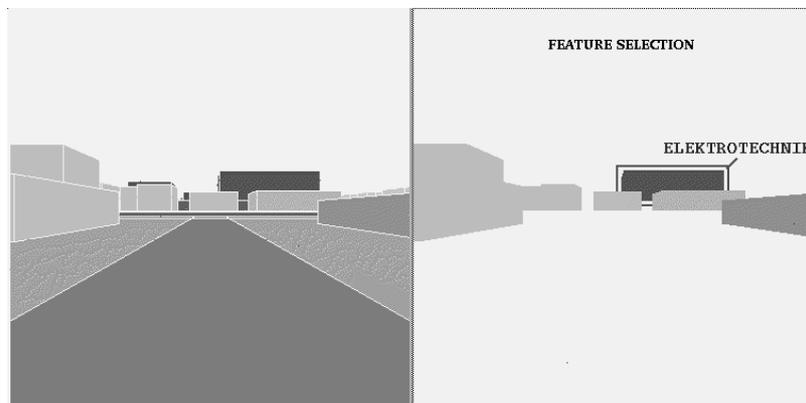


Abbildung 4.33: Beispiel der Auswahl eines salienten Objektes bei zentraler Fokussierung

#### 4.2.3.9 Adaption der visuellen Salienzen an den räumlichen Fokusbereich

In den bisherigen Funktionen ist der räumliche Fokus nur in soweit eingegangen, daß Lokationen bzw. Objekte innerhalb des Fokus zuerst betrachtet werden. Neben dieser Präferenzierung von Objekten bewirkt der visuelle Fokus zusätzlich eine Modifikation der visuellen Salienzwerte *vis-sal*. Visuelle Salienzen außerhalb der Fokus werden entsprechend ihrer Distanz vom Fokusbereich vermindert. Solche innerhalb des Fokus bleiben unverändert. Die Modellierung erfolgt über eine stetig abfallende Splinefunktion *focus* (vgl. Gleichung 19). Die visuelle Salienz eines Objektes  $vs(O)$  ergibt sich folglich aus der Evaluation der Funktion *focus* auf dem Ergebnis der Funktion *vis-sal*( $O$ ).

$$(19) \quad vs(O) = focus(vis - sal(O))$$

<sup>43</sup>Ein interessanter Ansatzpunkt ist es, Koeffizienten für verschiedene Domänen zu bestimmen und miteinander zu vergleichen.

#### 4.2.3.10 Zusammenfassung

In diesem Abschnitt ist ein Modell für die Bestimmung visueller Salienzen auf der Basis von visuellen Merkmalen vorgestellt worden. Die visuelle Salienz wird als Konzept verstanden, um Objekte aus einer visuell wahrnehmbaren Situation effizient nach Auffälligkeit auszuwählen. Bedingt durch die Bewegung des Agenten und den kommunikativen Kontext, ist eine räumliche Fokussierungsmöglichkeit integriert. Objekte im Fokus werden vor solchen betrachtet, die außerhalb liegen. Dies reduziert den Berechnungsaufwand selbst in komplexen Situationen mit einer Vielzahl von Objekten auf die Objekte, die im Fokus liegen. Eine Erweiterung dieses Konzeptes verändert den Öffnungswinkel des räumlichen Fokusbereichs entsprechend den zeitlichen Restriktionen.

Das Modell zur Berechnung visueller Salienzen basiert auf der Annahme, daß sich visuelle Daten in Merkmale zerlegen lassen (vgl. Kapitel 2). Merkmale lassen sich generell in rekursive und lineare Merkmale einteilen. Am Beispiel von Farbe ist gezeigt worden, wie die visuelle Salienz eines rekursiven Merkmals bestimmt wird. Durch Zerlegung einer Sicht mittels eines Rasters wird eine objektunabhängige Repräsentation dazu verwendet, einzelnen Lokationen visuelle Salienzen zuzuordnen. Für das Merkmal Farbe wird eine an die Farbwahrnehmung des Menschen approximativ adaptierte RGB-Zerlegung verwendet. Die visuelle Salienz eines Farbmerkmals ist einerseits durch die Verteilung der Werte des Farbvektors selbst und durch die Verteilung in der gesamten Situation gegeben. Als Beispiele für lineare Merkmale sind *Höhe* und *Breite* diskutiert worden. Die für die einzelnen Merkmale bestimmten visuellen Salienzen werden lokationserhaltend überlagert und mittels einer Linearkombination integriert. Abschließend werden diese visuellen Salienzen von Objekten an den fokussierten Bereich derart adaptiert, daß visuelle Salienzen, an die Objekte außerhalb des Fokus gebunden sind, vermindert werden. Letztendlich wird eine partielle Ordnung von Objekten entsprechend ihren visuellen Salienzwerten determiniert.

Im folgenden Abschnitt wird vorgestellt, wie die durch die Pfadauswahl und die Objektauswahl selektierte Information in räumliches Wissen transformiert wird.

#### 4.2.4 Funktionen des Raumintegrationsmoduls

Die Sequenz der Pfadelemente der aktuellen Situation und die nach visueller Salienz geordneten Objekte werden im *Raumintegrationsmodul* zueinander und zur Lokation des Agenten in räumliche Beziehung gesetzt. Referenzsysteme sind das strukturierende Konzept, mit dessen Hilfe räumliches Wissen etabliert und als Grundlage für die Bestimmung von sprachlichen Verhaltensweisen verwendet werden. Da Referenzsysteme für die Etablierung raumbezogenen Wissens primär sind, werden zuerst die in dynamischen Situationen verwendeten Referenzsysteme diskutiert (vgl. Kapitel 4.2.4.1). Auf der Grundlage dieser Referenzsysteme werden Objekte und Pfade zueinander und zum Agenten durch *räumliche Relationen* in Beziehung gesetzt (vgl. Kapitel 4.2.4.2). Integriert wird das raumbezogene Wissen in sogenannten räumlichen Konfigurationsbeschreibungen, die dynamisch erweitert werden. Diese Dynamik kommt bei der ressourcenadaptiven Sprachgenerierung von inkrementellen Wegbeschreibungen zum Tragen (vgl. Kapitel 4.2.6).

##### 4.2.4.1 Egozentrische und allozentrische räumliche Referenzsysteme

In *statischen Situationen* ruht sowohl die Orientierung und Lokation des Betrachters, als auch die der Objekte. Die meisten Arbeiten im Kontext der Raumkognition nehmen solche statischen Situationen an (vgl. Kapitel 2). In diesem Fall hat der Betrachter die Möglichkeit, Objekte durch Lagebeziehungen auf sein statisches Egozentrum zu beziehen oder aber auf eine mental eingenommene Lokation unter Einnahme einer Orientierung, die durch Eigenschaften der Lokation

oder des Objektes *intrinsisch* oder durch den Akteur *extrinsisch* vorgegeben sind (vgl. Kapitel 2.3.1, 2.5). Bewegt sich der Betrachter relativ zu seiner Umgebung, ändern sich einerseits seine Lagebeziehungen zu Objekten und andererseits die räumliche Verankerung der Referenzsysteme und deren Orientierung dynamisch. Hierzu wird die Verwendung von Referenzsystemen auf dynamische Situationen erweitert.

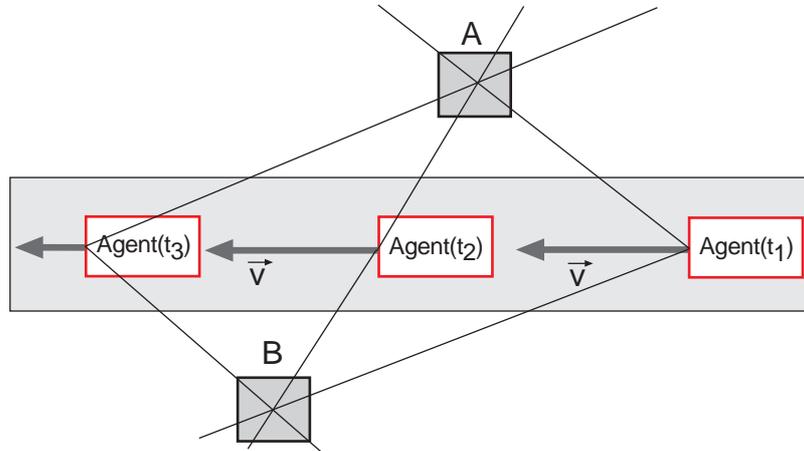


Abbildung 4.34: Bewegung eines Agenten auf einem Pfad in einer Umgebung mit zwei Landmarken  $A$  und  $B$  zu den Zeitpunkten  $t_1$ ,  $t_2$  und  $t_3$ .

Bewegt sich ein Agent durch eine Umgebung, so ändert sich die Ausrichtung und Verankerung des primären Referenzsystems und dadurch die räumlichen Lagebeziehungen von Objekten fortlaufend (vgl. Abbildung 4.34). Selbst wenn raumbezogenes Wissen nur in diskreten zeitlichen Abständen etabliert wird, muß es dennoch dynamisch an die Situation angepaßt werden. Die Lagebeziehungen zwischen den Landmarken  $A$  und  $B$  ändern sich permanent (indiziert durch die Verbindungslinien zum Agenten), bis sie aus dem Blickwinkel des Betrachters verschwinden und dadurch für das raumbezogene Wissen bezüglich einer Situation keine Rolle mehr spielen.

Ein *räumliches Referenzsystem* definiert die Orientierung des Agenten und die Strukturierung des Raums bezüglich einer Situation (vgl. Kapitel 2.3.1). Auf Grund der Struktur der Umgebung und raumsprachlicher Erfordernisse, insbesondere hinsichtlich räumlicher Präpositionen, basieren die im weiteren vorgestellten räumlichen Referenzsysteme auf kartesischen Dimensionszerlegungen. Weiterhin wird für die Generierung von intrinsischen und extrinsischen Beschreibungen angenommen, daß der Agent in der Lage ist, seinen Betrachterstandpunkt mental zu verschieben (vgl. Kapitel 2.3). Hierfür ist der Aufbau von Referenzsystemen notwendig, welche *virtuell* in Lokationen außerhalb des Agenten verankert sind. Die Annahme, daß eine beliebige Orientierung und Verankerung eines Referenzsystems möglich ist, wird in dieser Arbeit dahingehend verfeinert, daß die verwendeten Referenzsysteme von einem egozentrischen Referenzsystem abgeleitet werden. Eine solche Hypothese ist durch die Eigenbewegung des Agenten bedingt. Einerseits ist durch die Bewegungsrichtung ein Referenzsystem vorgegeben. Andererseits kann ein Akteur zusätzlich durch Fokussierung von räumlichen Konfigurationen neben dieser Hauptrichtung andere Referenzsysteme etablieren und sich imaginär in Objekte oder Lokationen hineinversetzen. Eine solche Änderung ist vom kognitiven Aufwand her durch die mentalen Operationen *Rotation* und *Translation* bestimmt. Dies stützt sich u.a. auf Ergebnisse der Experimente von Herrmann et al. (vgl. Kapitel 2.3.1, insbesondere [HGH91]). Danach sind Reaktionszeiten im sogenannten *Elastizitätsbereich* ( $-60^\circ$  bis  $+60^\circ$ ) einer egozentrischen Orientierung minimal. Bei Verwendung von allozentrischen Referenzsystemen steigen die Reaktionszeiten streng monoton proportional zum Rotationswinkel an (vgl. Kapitel 2.3.1). Nach Herrman et al. nehmen Re-

aktionszeiten bei egozentrierten Referenzsystemen nach einem Anstieg mit Maximum bei  $120^\circ$  wieder ab. Hieraus folgt grundlegend, daß Rotationen von Referenzsystemen beim „Sichhineinversetzen“ mit kognitivem Mehraufwand verbunden sind. Aufgrund kognitiver Ökonomie folgt ebenfalls eine Bevorzugung egozentrischer Referenzsysteme, die sich auch empirisch nachweisen läßt (vgl. Kapitel 2.3.1, 2.5). Im Elastizitätsbereich hingegen, kann die Orientierung egozentrischer Referenzsysteme nahezu beliebig mental rotiert werden.

Da dies besonders in dynamischen Situationen von Bedeutung ist, werden zwei Klassen von egozentrischen Referenzsystemen eingeführt. Die erste Klasse der *primären Referenzsysteme* (vgl. Kapitel 4.2.4.1.1) entspricht dem in der Literatur indizierten egozentrischen Referenzsystem (vgl. Kapitel 2.3.1). Der Agent unterhält permanent ein primäres Referenzsystem (PRS), welches durch die Orientierung der *Bewegungs-*, *Aktions-* oder *Perzeptionsrichtung* festgelegt ist. Eine zweite Klasse ergibt sich, wenn die Hauptorientierung des Betrachters sich kurzfristig ändert, jedoch im Elastizitätsbereich verbleibt. Referenzsysteme dieser Klasse werden als *sekundär* bezeichnet (vgl. Kapitel 4.2.4.1.2). Allozentrische Referenzsysteme werden entsprechend empirischen und konzeptuellen Arbeiten in den betrachteten Objekten verankert (vgl. Kapitel 2.3.1 und z.B. [Her86]). Ein flexiblerer Ansatz ist es anzunehmen, daß sich allozentrische Referenzsysteme in Lokationen verankern lassen. Dies schließt den Fall mit ein, daß an einer Lokation sich ein Objekt befindet, welches durch intrinsische Eigenschaften die Orientierung des Referenzsystems modifiziert. Ebenso umfaßt ein solches virtuelles Referenzsystemkonzept solche Referenzsysteme, die bezüglich einer Überblicks- oder Umgebungsperspektive eingenommen werden (z.B. [FTC92]). Als eine Randbedingung läßt sich auf Grund kognitiver Ökonomie sowie aus obigen Befunden ableiten, daß die Orientierung eines virtuellen Referenzsystems nach Möglichkeit durch eine minimale mentale Rotation bezüglich des primären Referenzsystems erreicht wird.

Es werden im folgenden drei Referenzsystemklassen (primär, sekundär und virtuell) diskutiert. Dabei wird gezeigt, wie sie dazu verwendet werden, raumbezogenes Wissen in dynamischen Umgebungen zu etablieren.

**4.2.4.1.1 Das primäre Referenzsystem** Das *primäre Referenzsystem* (PRS(x, y, z)) basiert auf dem Konzept eines kartesischen Koordinatensystems mit drei unabhängigen Dimensionen, die den dreidimensionalen Raum aufspannen (vgl. Kapitel 2.3.1). Die Ausrichtung der z-Achse (oben/unten) ist parallel zur Gravitationskraft. Die Orientierung der Frontalrichtung des primären Referenzsystems ist ausschließlich durch die Hauptachse festgelegt, die durch die Richtung der *Bewegung*, der *Perzeption* oder der *Aktivität* des Agenten determiniert ist (vgl. Abbildung 4.35). Die Ausrichtung in der Transversalen leitet sich von der Frontalrichtung ab. Bewegt sich der Agent auf einem Pfad, so ist die Hauptachse durch die Trajektorie der Bewegung determiniert (vgl. Abbildung 4.36). Die Hauptperzeptionsrichtung ist im Falle einer Vorwärtsbewegung mit der Frontalrichtung meist identisch, so daß sich das primäre Referenzsystem wesentlich von der Hauptausrichtung der Bewegung und der Perzeption ableitet.

In dynamischen Situationen, in denen der Betrachter über verschiedene Freiheitsgrade verfügt, kann die perzeptuelle Hauptrichtung kurzfristig modifiziert werden, indem der Agent seine Aufmerksamkeit auf ein bestimmtes Gebiet im Sichtfeld fokussiert. Ein solcher Fall liegt außerhalb der Eigenschaften primärer Referenzsysteme, weswegen es einer zusätzlichen Klasse von Referenzsystemen bedarf.

**4.2.4.1.2 Das sekundäre Referenzsystem** Bei kurzfristigen Änderungen der Hauptausrichtung, besonders der Perzeption, verfügt der Agent über ein *sekundäres Referenzsystem* (SRS), welches es ihm ermöglicht, räumliche Sachverhalte lokal räumlich zu erfassen. Nachdem dieser räumliche Sachverhalt nicht weiter fokussiert wird, wird das sekundäre Referenzsystem

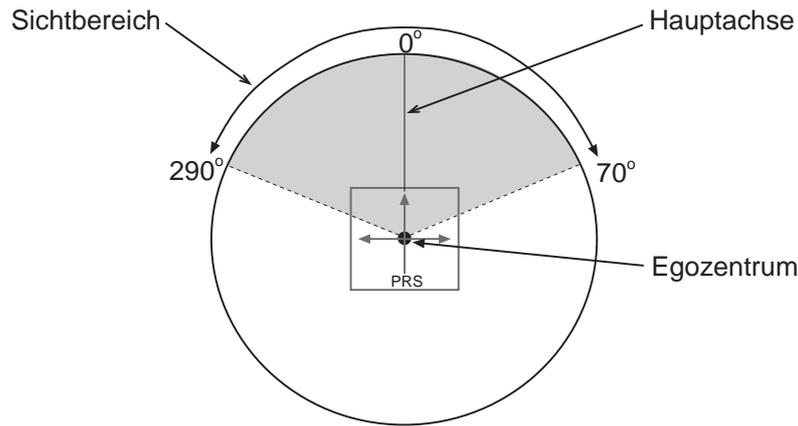


Abbildung 4.35: Orientierung des primären Referenzsystems in der horizontalen Ebene erfolgt immer in Ausrichtung der Hauptachse (entlang Winkel  $\alpha = 0^\circ$ )

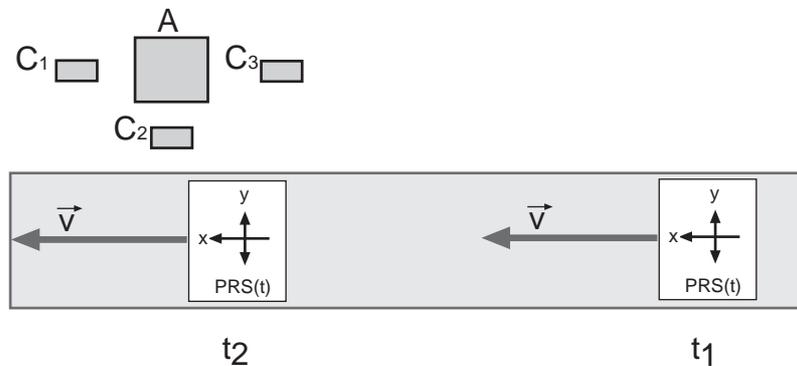


Abbildung 4.36: Primäres Referenzsystem zum Zeitpunkt  $t$ . Der Agent bewegt sich entlang der Hauptachse des Pfades mit einer Geschwindigkeit  $\vec{v}$  an einer räumlichen Konfiguration mit den Objekten  $A$ ,  $C_1$ ,  $C_2$  und  $C_3$  vorbei

eliminiert und wieder das primäre Referenzsystem eingenommen. Die Notwendigkeit eines *SRS* stellt sich ausschließlich in Umgebungen, in denen der Agent den Freiheitsgrad, der durch perzeptuelle Aufmerksamkeitssteuerungen entsteht, besitzt. Ebenso wie das primäre Referenzsystem, ist ein sekundäres Referenzsystem im Egozentrum des Betrachters verankert (vgl. Abbildung 4.37). Formal wird ein *SRS* von einem *PRS* durch eine Rotation um den Winkel  $\alpha$  abgeleitet:  $ROT_\alpha$ .

$$(20) \quad SRS(x, y, z) = ROT_\alpha(PRS(x, y, z))$$

Ist ein *SRS* etabliert, so wird dieses bevorzugt verwendet. Deiktische sprachliche Beschreibungen erfolgen in solchen Fällen ausschließlich bezüglich dieses *SRS*. In Abbildung 4.38 bewegt sich der Agent beispielsweise entlang eines Pfades und perzipiert zum Zeitpunkt  $t_1$  einen räumlichen Sachverhalt, der durch die Landmarken  $A$ ,  $C_1$ ,  $C_2$  und  $C_3$  gegeben ist. Es wird angenommen, daß die Ausrichtung der Aufmerksamkeit in Richtung des Schwerpunkts des räumlichen Sachverhalts durch Rotation um den Winkel  $\alpha$  erfolgt. Entsprechend diesem Rotationswinkel wird ein *SRS*

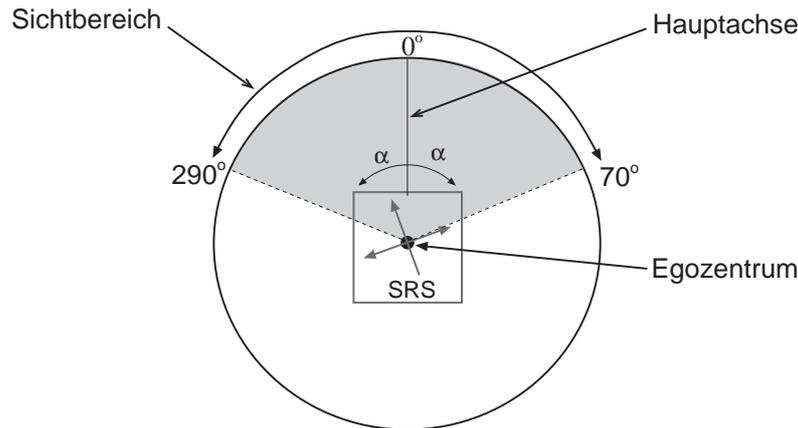


Abbildung 4.37: Orientierung des sekundären Referenzsystems

zum Zeitpunkt  $t_1$  etabliert. Bewegt sich der Agent mit einer Geschwindigkeit von  $\vec{v}$  entlang des Pfades, so erreicht er zu einem Zeitpunkt  $t_2$  eine Lokation, bei der der Rotationswinkel  $\alpha$  nahezu  $90^\circ$  beträgt. Zum Zeitpunkt  $t_1$  sind PRS und SRS in ihrer Orientierung nahezu identisch. Diese Übereinstimmung verstärkt die Verwendung dieser beiden Referenzsysteme zur Beschreibung der Situation (vgl. [CRI93]). Aus diesem Grund wird die Situation zum Zeitpunkt  $t_1$  durch Verwendung des sekundären Referenzsystems beschrieben („ $C_3$  ist vor A.“). Zum Zeitpunkt  $t_2$  ändert sich der Zusammenhang zwischen SRS und PRS. Beide verstärken sich nicht mehr, so daß zwischen beiden gewählt werden muß.<sup>44</sup> Um in dieser Situation  $C_3$  noch visuell wahrnehmen zu können, muß SRS eingenommen werden. Von PRS aus ist  $C_3$  zum Zeitpunkt  $t_2$  nicht mehr zugreifbar, so daß eine entsprechende Beschreibung unangebracht ist (vgl. Satz 21). Somit wird in dieser Situation das SRS ausgewählt.

$$(21) \quad \begin{array}{l} \text{PRS } t_1: \text{ „}C_3 \text{ ist vor A.“} \\ \text{PRS } t_2: \text{ —} \\ \text{SRS } t_1: \text{ „}C_3 \text{ ist vor A.“} \\ \text{SRS } t_2: \text{ ? „}C_3 \text{ ist rechts von A.“} \end{array}$$

Ausgehend von einer beschränkten Rotationsmöglichkeit des perzeptuellen Apparates, ist nur eine Rotation des *SRS* im Winkelbereich zwischen  $190^\circ$  und  $70^\circ$  zulässig. Es wird angenommen, daß für Fälle, in denen der Agent seine Aufmerksamkeit über diesen Winkelbereich hinaus rotieren möchte, er gleichfalls seine Hauptachse, d.h. die Frontalausrichtung des *PRS* ändert (vgl. Abbildung 4.39). Mit der Etablierung eines neuen *PRS* wird gleichzeitig das *SRS* entfernt. In solchen Fällen führt dies zu einer erheblichen Veränderung der Hauptachse, welches gleichbedeutend mit der Konstituierung eines *Verzweigungspunktes* ist.

Wie extrinsische und intrinsische Beschreibungen indizieren, ist ein Sprecher in der Lage, auch räumliche Beziehungen relativ zu einem Referenzsystem außerhalb seines Egozentrum zu etablieren. Welche Referenzsysteme hierzu notwendig sind, wird im folgenden vorgestellt.

**4.2.4.1.3 Das virtuelle Referenzsystem** Allozentrische Referenzsysteme gehören nicht zur Klasse *primärer* und *sekundärer Referenzsysteme*, weswegen eine dritte Klasse von Referenzsystemen erforderlich ist, die *virtuell* in einer beliebigen Lokation einer Situation verankert

<sup>44</sup>Nicht betrachtet wird hierbei die Möglichkeit, den Rezipienten explizit darauf hinzuweisen, ein sekundäres Referenzsystem zu verwenden.

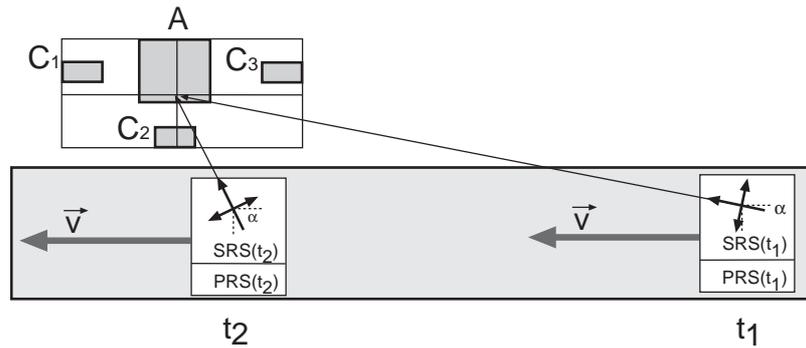


Abbildung 4.38: Sekundäres Referenzsystems zum Zeitpunkt  $t$ . Ausrichtung des SRS auf den Schwerpunkt der Konfiguration zu den Zeitpunkten  $t_1$  und  $t_2$  bei einer Bewegung mit Geschwindigkeit  $\vec{v}$ .

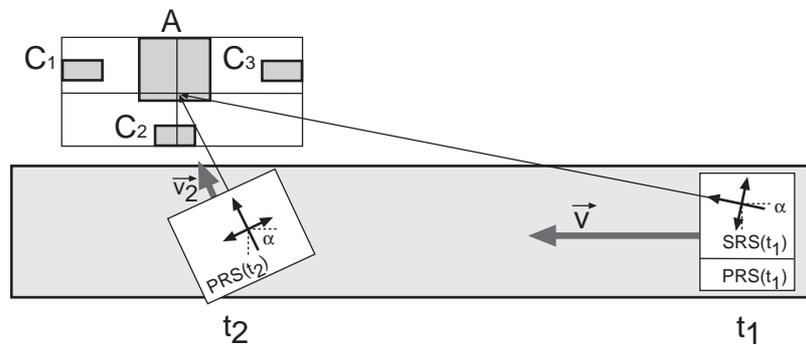


Abbildung 4.39: Änderung der Hauptachse des  $PRS$  zu einem Zeitpunkt  $t_2$ .

sind (VRS). Im Gegensatz zu einigen raumsprachlichen Arbeiten (vgl. [Cla73], [Her86] und Kapitel 2.5, 2.3.1) wird davon ausgegangen, daß die Etablierung von allozentrischen Referenzsystemen mit unterschiedlichem kognitiven Aufwand verbunden ist. Besonders deutlich wird im Experiment von Herrmann et al. (vgl. Kapitel 2.3.1 und [HGH91]) die Abhängigkeit von Rotationswinkeln beim „Sichhineinversetzen“ in eine Lokation der Situation. Entsprechend dieser Experimentalreihe lassen sich Beschränkungen an virtuellen Referenzsystemen (VRS) definieren. Durch imaginierte Translation wird ein VRS in einer Lokation der Situation außerhalb des Agenten verankert. Es werden solche Rotationswinkel bevorzugt, die sich im *Elastizitätsbereich*, d.h. einer Winkelabweichung  $\alpha$  von  $60^\circ$  relativ zur Hauptachse, befinden. Dabei wird angenommen, daß das VRS durch eine imaginierte Rotation aus dem  $PRS$  bzw., falls etabliert, aus dem  $SRS$ , entsteht (vgl. Abbildung 4.40). Formal entspricht dies einer Rotation des  $PRS$  bzw.  $SRS$  um den Winkel  $\alpha$ :  $ROT_\alpha$ .

$$(22) \quad \text{VRS}(x, y, z) = \text{ROT}_\alpha(\text{TRANS}_s(\text{SRS}(x, y, z))), \\ \text{mit } 300^\circ \leq \alpha \leq 60^\circ$$

Ein VRS ist, ebenso wie das SRS, lokal bezüglich einer Situation und einer Aufmerksamkeitsausrichtung definiert (siehe Abbildung 4.41). Dabei werden solche VRS bevorzugt, die parallel zu einer der Hauptachsen ( $h_1$  und  $h_2$  in Abbildung 4.41) des fokussierten räumlichen Sachverhalts liegen (vgl.  $\text{VRS}_1$  und  $\text{VRS}_2$  in Abbildung 4.41).

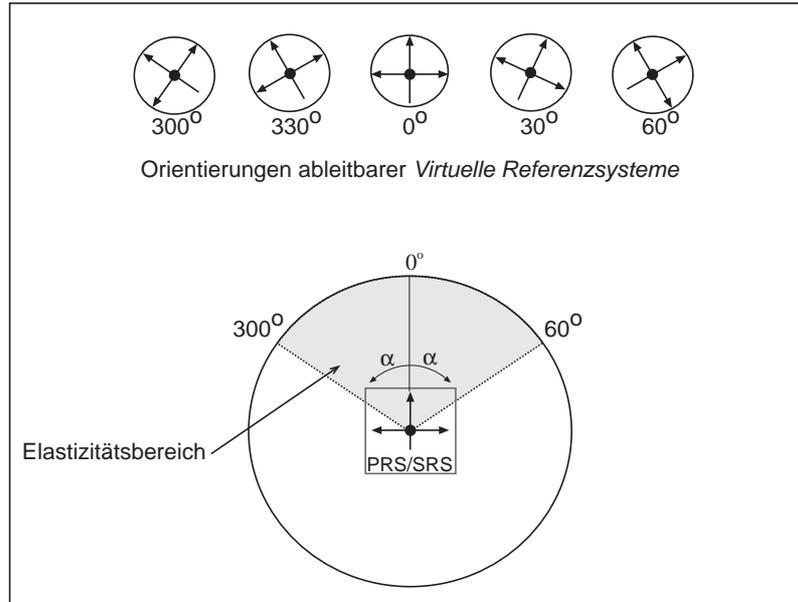


Abbildung 4.40: Orientierung eines virtuellen Referenzsystems

Zwei Fälle sind bei der Etablierung eines virtuellen Referenzsystems zu unterscheiden.

1. *Extrinsische* Etablierung eines virtuellen Referenzsystems
2. *Intrinsische* Etablierung eines virtuellen Referenzsystems

Im ersten Falle wird das Referenzsystem außerhalb und im zweiten Fall innerhalb eines Objekts verankert. Die Behandlung einer intrinsischen Etablierung berücksichtigt intrinsische Eigenschaften des Referenzobjekts. Dieser Fall wird in Abschnitt 4.2.4.1.6 diskutiert. Die extrinsische Etablierung eines virtuellen Referenzsystems wird im folgenden betrachtet (vgl. Abschnitt 4.2.4.1.4, 4.2.4.1.5).

**4.2.4.1.4 Etablierung von virtuellen Referenzsystemen bei orthogonalen Konfigurationen** Bei einer extrinsischen Etablierung wird das virtuelle Referenzsystem in einer Lokation unmittelbar in der Nähe des fokussierten räumlichen Sachverhalts verankert. In Abbildung 4.41 befindet sich das virtuelle Referenzsystem  $VRS_1(t)$  zum Zeitpunkt  $t$  bezüglich der Hauptachse  $h_1$  an Lokation  $l_1$  bei einer Rotation  $\beta$  relativ zur Orientierung des sekundären Referenzsystems  $SRS(t)$ . Gleichzeitig kann ein  $VRS_2(t)$  bezüglich Hauptachse  $h_2$  an Lokation  $l_2$  etabliert werden, was jedoch aufgrund des großen Winkels  $\alpha$  eine geringe Anwendbarkeit besitzt.

**4.2.4.1.5 Etablierung von virtuellen Referenzsystemen bei nicht-orthogonalen Konfigurationen** Bisher ist von Situationen ausgegangen worden, deren Lage sich parallel oder orthogonal zu der Bewegungsrichtung des Agenten fand. Ist dies nicht der Fall, so stehen die Hauptachse des Agenten und die Hauptachsen,  $h_1$  und  $h_2$ , des fokussierten Sachverhalts weder in einer parallelen noch einer orthogonalen Lagebeziehung (vgl. Abbildung 4.42).

Das PRS ist wie bisher durch die Hauptbewegungsrichtung entlang des Pfades gegeben. Das SRS leitet sich vom Rotationswinkel  $\beta$  zwischen der Frontalachse des PRS und dem Schwerpunkt des räumlichen Sachverhalts (Landmarken  $A$  und  $C$ ) ab (vgl. Abbildung 4.42). Potentiell sind drei *virtuelle Referenzsysteme*  $VRS_1$ ,  $VRS_2$  und  $VRS_3$  etablierbar.  $VRS_2$  und  $VRS_3$  werden

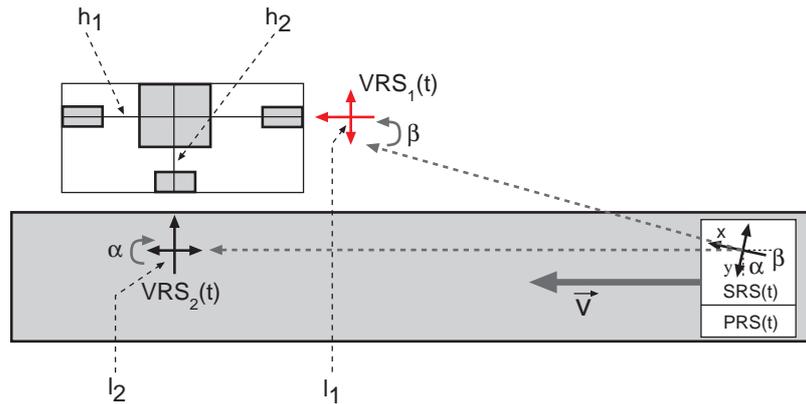


Abbildung 4.41: Virtuelle Referenzsysteme ( $VRS_1(t)$ ,  $VRS_2(t)$ ) zum Zeitpunkt  $t$  bezüglich der Hauptachsen  $h_1$  und  $h_2$ .

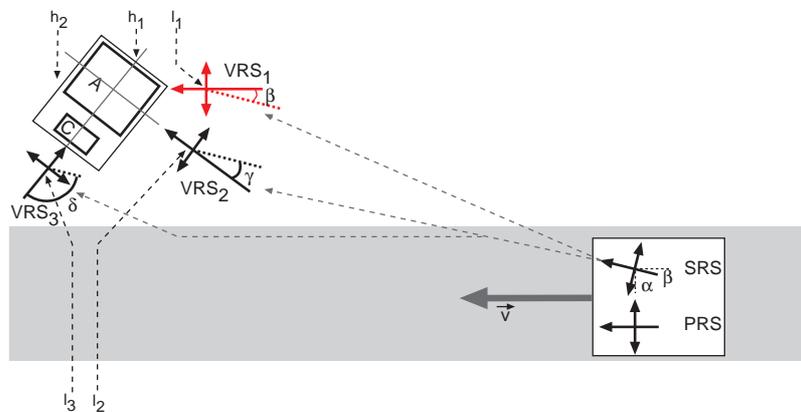


Abbildung 4.42: Virtuelle Referenzsysteme bei geneigten Konfigurationen

bezüglich der beiden Hauptachsen  $h_1$  und  $h_2$  des Referenzobjektes  $A$  an den Lokationen  $l_2$  und  $l_3$  definiert. Da zwischen dem Agenten und dem räumlichen Sachverhalt eine große Distanz besteht, sind die beiden Winkel  $\delta$  und  $\gamma$  groß. In diesem Fall bietet sich aus kognitiv-ökonomischen Gründen eine alternative Verankerungsmöglichkeit durch Rückgriff auf das *primäre Referenzsystem* an.  $VRS_1$  wird in einer Lokation  $l_1$  verankert, die sich auf einer durch den Schwerpunkt des Referenzobjektes gehenden Linie parallel zur Bewegungsrichtung befindet. Es lassen sich somit für die Verankerung eines geeigneten virtuellen Referenzsystems im allgemeinen zwei Fälle unterscheiden.

1. Verankerung bezüglich der Hauptachsen des Referenzsystems
2. Verankerung bezüglich der Hauptachse des primären Referenzsystems

Nicht in jedem Fall kann vom Referenzobjekt in dieser Weise abstrahiert werden, daß die Information über den Schwerpunkt und die Hauptachsen ausreichen, was im folgenden behandelt wird.

**4.2.4.1.6 Berücksichtigung intrinsischer Eigenschaften bei der Wahl eines virtuellen Referenzsystems** In vielen Fällen besitzt das Referenzobjekt ausgezeichnete intrinsische

Eigenschaften, die die Auswahl eines geeigneten virtuellen Referenzsystems beeinflussen (vgl. Kapitel 2.3.1). Nimmt man beispielsweise an, daß das Referenzobjekt  $A$  in Abbildung 4.42 eine Kirche ist, dann ist die Front der Kirche durch ihr Portal gegeben. Die Objektklasse der Kirchen besitzt die intrinsische Eigenschaft, ein virtuelles Referenzsystem zu prädeternieren. Wie beispielsweise nachgewiesen werden konnte, wird in der deutschen Sprache bezüglich einer Kirche ein virtuelles Referenzsystem entsprechend  $VRS_3$  bevorzugt (vgl. [Car93]).<sup>45</sup> Ein Betrachter nimmt somit eine Perspektive ein, als ob er von außen auf das Portal der Kirche schauen würde. Solche intrinsischen Eigenschaften sind funktionale, soziale, kulturelle oder ähnliche Merkmale, die einem einzelnen oder einer Klasse von Objekten zugeordnet werden.

Ein Spezialfall tritt auf, wenn der Betrachter durch häufigen Umgang oder durch Konventionen annimmt, daß er sich an einer prototypischen Lokation unter Einnahme einer prototypischen Orientierung im Referenzobjekt selbst befindet. Ein Beispiel hierfür ist das Automobil. Durch häufigen Gebrauch bedingt, versetzt sich ein Betrachter typischerweise in die Position des Fahrers oder Beifahrers mit Orientierung in Fahrtrichtung. Eine solche prototypisch vorgegebene Verankerung bestimmt die räumliche Verankerung und Orientierung eines intrinsisch determinierten, virtuellen Referenzsystems  $VRS_{int}$  (vgl. Abbildung 4.43). Anders ist dies für Referenzobjekte, mit denen ein Betrachter sich i.a. nicht identifizieren kann, wie beispielsweise mit eines Nagels in der Wand. Lokalisiert man ein Bild an der Wand bezüglich einem Nagel, so nimmt man typischerweise nicht die Position des Nagels ein, obschon ein Nagel prädominante intrinsische Eigenschaften besitzt.<sup>46</sup>

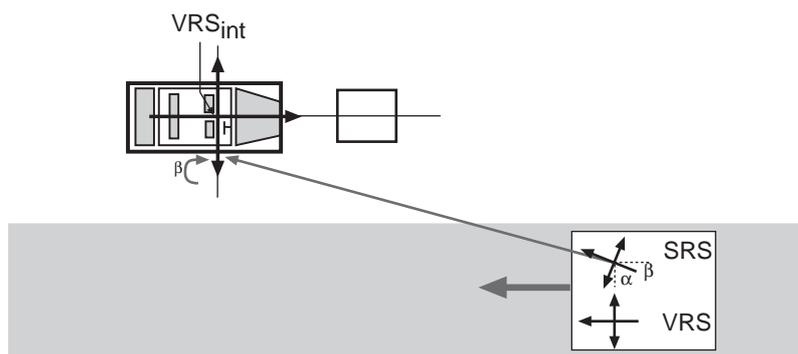


Abbildung 4.43: Verankerung eines intrinsisch determinierten, virtuellen Referenzsystems ( $VRS_{int}$ ) in einem Referenzobjekt (Automobil) aufgrund prototypischen Wissens über den Umgang mit diesem.

Intrinsisch determinierte virtuelle Referenzsysteme sind weitgehend von Rotationswinkeln unabhängig. Aufgrund der Untersuchungen von Herrmann et al. ist zu vermuten, daß dies daher rührt, daß der kommunikative Akt einer Raumbeschreibung durch ein solches prototypisches Wissen und durch Konventionen erheblich vereinfacht wird. Eine solche Erklärung läßt sich auch auf Adressaten einer Beschreibung ausweiten. Wie in mehreren Untersuchungen gezeigt worden ist, werden Beschreibungen in kommunikativen Kontexten präferiert, in denen der Sprecher die Position des Rezipienten einnimmt (vgl. [HBN87], [Sch95b], [Som69]).

<sup>45</sup>Interessanterweise etablieren englischsprachliche Personen das virtuelle Referenzsystem so, als ob sie aus dem Portal hinausschauen würden. Dadurch bedingt ist eine Spiegelung dessen, was als „links“ von und „rechts“ von der Kirche bezeichnet wird.

<sup>46</sup>Interessant sind Fälle, bei denen man nicht allgemein für eine Objektklasse bestimmen kann, ob das  $VRS$  extern oder intern zum Referenzobjekt verankert wird. Als Beispiel sei die Klasse der Gebäude genannt. Spricht man über das Elternhaus, so ist das  $VRS$  typischerweise intern verankert. Für den Deutschen Reichstag hingegen wird typischerweise eine externe Verankerung verwendet.

Zusammenfassend lassen sich mit *primären*, *sekundären* und *virtuellen* drei verschiedene Klassen von Referenzsystemen identifizieren, die in dynamischen Situationen grundlegend für die Verankerung raumbezogenen Wissens sind. Primäre und sekundäre Referenzsysteme sind egozentrisch, wohingegen virtuelle Referenzsysteme allozentrisch sind. Zur Auswahl eines geeigneten virtuellen Referenzsystems wird eine extrinsische und eine intrinsische Strategie verwendet. Die extrinsische bezieht sich einzig auf geometrische Eigenschaften des Referenzobjektes, wie Hauptachsen und Schwerpunkte. Die intrinsische Strategie basiert auf intrinsischen Eigenschaften des betrachteten Objektes, die aus verschiedenen Quellen stammen können und deren Formalisierung i.a. komplex ist.

#### 4.2.4.2 Evaluation räumlicher Relationen

In vorhergehenden Abschnitten ist beschrieben worden, wie verschiedene Referenzsysteme in dynamischen Situationen etabliert werden. Bezüglich eines Referenzsystems lassen sich Objekte durch räumliche Lagebeziehungen, d.h. *räumliche Relationen*, zueinander und zum Betrachter zu einer *räumlichen Vorstellung* integrieren. Es wird zwischen *topologischen* und *projektiven* (deiktischen und intrinsischen) räumlichen Relationen unterschieden (vgl. Kapitel 2.3.2). Deiktische und intrinsische räumliche Relationen basieren auf dem Konzept des Referenzsystems. Als einzige sind topologische räumliche Relationen, entsprechend ihrer Definition, von Referenzsystemen unabhängig. Sie repräsentieren qualitative Lagebeziehungen zwischen Objekten. Durch die explizite Berücksichtigung der Kontaktbeziehung repräsentieren topologische räumliche Relationen vor allem physische Beziehungen.

Bei der Evaluation von räumlichen Relationen lassen sich mehrere Ebenen unterscheiden. In den meisten Arbeiten wird eine Unterteilung in zwei Ebenen angenommen, die das Konzept eines Objektes in geometrisch-strukturelle und semantische Eigenschaften zerlegen (vgl. Kapitel 2.3.2). Die Evaluation räumlicher Relationen, wie die auch hier verwendete, basiert nahezu ausschließlich auf einer geometrisch-strukturellen Repräsentationsebene. Semantische Eigenschaften von Objekten spielen nur eine eingeschränkte Rolle (vgl. [Her86], [LCS90], [Gap94]).

In diesem Kapitel werden die drei verschiedenen Relationenarten (topologisch, deiktisch und intrinsisch) nacheinander soweit diskutiert, wie sie für die Verwendung in dieser Arbeit notwendig sind.

##### 4.2.4.2.1 Prinzip zur Bestimmung von Anwendbarkeitsgraden über Lokationen

Räumliche Relationen sind im sprachlichen Gebrauch nicht durch zweiwertige Logikprädikate definierbar. Vielmehr sind räumliche Relationen unscharfe Konzepte, deren Verwendung graduelle Unterschiede aufweisen (z.B. [Her86], [Han80]). Auf sprachlicher Ebene sind linguistische Hecken Markierungen solcher Vagheiten (vgl. [Lak72]). Zur Formalisierung werden die Gebiete um das Referenzobjekt *ro* in diskrete räumliche Lokationen<sup>47</sup> unterteilt (z.B. [Her96], [Gap94], [OMT94], [Hab87]). Ein Beispiel eines solche gebietskonstituierenden Ansatzes ist das in VITRA entwickelte räumliche Relationenmodell (vgl. [ABHR87], [RS88], [AHR89], [Gap94]). In diesem Modell wird einer jeden Relation über eine *Anwendbarkeitsfunktion* eine *Possibilität* zugeordnet, mit der diese Relation auf sprachlicher Ebene direkt als sprachliche Präposition verwendet wird (vgl. [SBSZ87]). Jeder Lokation ist dadurch bezüglich einer Anwendbarkeitsfunktion eine Gradierung auf dem Intervall  $[0...1]$  zugeordnet, die angibt, inwieweit eine assoziierte räumliche Präposition den betrachteten räumlichen Sachverhalt zwischen dem Referenzobjekt und dem Lokationsobjekt *lo* beschreibt. Als Eingabe gehen in diese Funktionen die Größe des Referenzobjektes sowie die Distanz und die Winkelabweichung des Lokalisationsobjektes zum Referenzobjekt mit ein (vgl. [Gap94]). Basierend auf der These von Landau und Jackendoff (vgl. [LJ93]) wird in

<sup>47</sup>Räumliche Lokationen sind dabei entweder zwei- oder null-dimensional.

diesem Modell davon ausgegangen, daß nur stark abstrahierte geometrische Eigenschaften des Referenzobjektes und des Lokationsobjektes in die Evaluation räumlicher Relationen eingehen. Das Referenzobjekt wird dabei durch einen minimal-umschreibenden Quader und *lo* durch seinen Schwerpunkt repräsentiert.<sup>48</sup> Die Größe des Referenzobjektes geht durch eine Normierung des Koordinatensystems ein (vgl. [Gap94]). Dabei wird angenommen, daß der Raum *homogen* ist, d.h., daß es keine ausgezeichneten Lokationen im Raum gibt und daß der Raum verzerrungsfrei bezüglich einer linearen Metrik ist.<sup>49</sup> Unter dieser Voraussetzung kann eine Situation bezüglich der Größe des Referenzobjektes normiert werden.<sup>50</sup> Einem solchen Ansatz liegt eine vollständige geometrische Modellierung einer Situation zugrunde, von der sich Distanzen und Winkelabweichungen direkt ableiten lassen. Inwieweit eine solche vollständige Repräsentation kognitive Plausibilität besitzt, ist eine offene Frage. In dem Modell werden ausschließlich statische Situationen betrachtet (z.B. [Gap94]). Eine Bezugnahme auf die Lokation des Sprechers wird dabei nur implizit verwendet, weswegen eine räumliche Relation  $R$  allgemein als binär zwischen dem Lokalisations- und Referenzobjekt definiert ist ( $R(lo, ro)$ , z.B. [MJL76]). Bewegt sich der Sprecher jedoch, ändern sich gleichzeitig über die Lokation des Agenten die egozentrischen und virtuellen Referenzsysteme. Über Referenzsysteme bestimmt sich, welche Art der Beschreibung (deiktisch, intrinsisch oder extrinsisch) verwendet wird. Dadurch geht die Bewegung des Agenten direkt in die Berechnung räumlicher Relationen mit ein, weswegen im weiteren räumliche Relationen als dreistellig verstanden werden:  $R(ego, lo, ro)$ . Nacheinander wird im weiteren diskutiert, in welcher Weise topologische, deiktische und intrinsische räumliche Relationen in dieser Arbeit verwendet werden. Für weitergehende Ausführungen zum Konzept der räumlichen Relationen sei auf Kapitel 2.3.2 und entsprechende Arbeiten verwiesen (vor allem [MJL76], [Her86], [Pri93], [Her96], [Gap94]).

**4.2.4.2.2 Topologische räumliche Relationen** Topologische räumliche Relationen sind unabhängig von Referenzsystemen. In der hier betrachteten Domäne der Wegbeschreibungen sind sie nur insoweit wichtig, daß sie Kontaktrelationen zwischen Pfadelementen repräsentieren. Um sicherzustellen, daß ein Agent direkt von einem Pfadelement zum nächsten gelangen kann, muß mindestens eine Kontaktrelation zwischen beiden bestehen. Legt man eine Punkt-mengenrepräsentation der Objekte zugrunde, so liegt zwischen zwei Objekten  $A$  und  $B$  dann eine Kontaktrelation vor, wenn wenigstens ein Punkt sowohl in  $A$ , als auch in  $B$  ist. In diesem Falle gilt die Relation  $\#meet\#(loc(A), loc(B))$ <sup>51</sup>. Eine solche Definition einer Kontaktrelation beruht auf einer Formalisierung von topologischen Relationen im zwei- und dreidimensionalen Raum (vgl. Kapitel 2.3.2, insb. [Ege91], [EAT92], [Rup96]). Danach lassen sich sechs topologische Relationen unterscheiden (vgl. Abbildung ??).

topo-rel]

Die Evaluation von topologischen Relationen zwischen zwei zweidimensionalen Objekten *obj<sub>1</sub>*

<sup>48</sup>Dies gilt aber nur unter der Annahme, daß die Objekte konvex sind.

<sup>49</sup>Es gibt empirische Zweifel an einer grundsätzlichen Homogenität. Beispielsweise zeigen Sadalla et al., daß erinnerte Distanzen zwischen einer Landmarke und einem Lokalisationsobjekt kleiner geschätzt werden als umgekehrt (vgl. [SBS80]). Optische Täuschungen, wie die von Müller-Lyer zeigen, daß es auch bei visuell Wahrgenommenem zu Verzerrungen kommen kann. Solche Phänomene lassen sich mit den vorgestellten Modellen bisher nicht formalisieren.

<sup>50</sup>Über diese Modellierung wird das Phänomen repräsentiert, daß sich bezüglich größerer Referenzobjekte räumliche Relationen über einen größeren räumlichen Bereich als anwendbarer erweisen als bei kleineren.

<sup>51</sup>Anstelle von  $loc(A)$  wird im weiteren abkürzend auf die explizite Nennung der Funktion  $loc$  verzichtet.

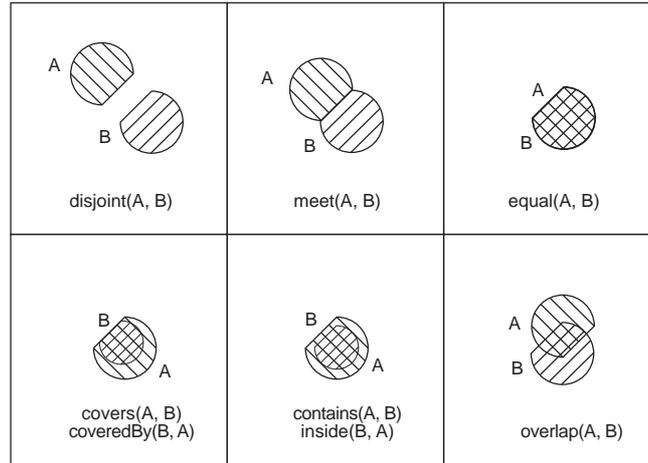


Abbildung 4.44: Topologische Relationen im zwei-dimensionalen Raum (vgl. [Ege91])

und  $obj_2$  erfolgt durch die Funktion *Evaluate-Topological-Rels* (vgl. Definition 23).<sup>52</sup>

$$(23) \quad \text{Evaluate-Topological-Rels}(\text{loc}(obj_1), \text{loc}(obj_2));$$

**4.2.4.2.3 Deiktische Relationen** Die Evaluation einer deiktischen Relation basiert entweder auf einem primären oder einem sekundären Referenzsystem, welches in der Lokation des Agenten (ego) verankert ist. Entsprechend der Hauptachsen eines kartesischen Referenzsystem werden sechs deiktische Relationen unterschieden (vgl. [Her86], [Log95], [HT95]):

1. #left-of#(ego, lo, ego):  $lo$  befindet sich links vom Agenten
2. #right-of#(ego, lo, ego):  $lo$  befindet sich rechts vom Agenten
3. #behind#(ego, lo, ego):  $lo$  befindet sich hinter dem Agenten
4. #in-front-of#(ego, lo, ego):  $lo$  befindet sich vor dem Agenten
5. #above#(ego, lo, ego):  $lo$  befindet sich über dem Agenten
6. #under#(ego, lo, ego):  $lo$  befindet sich unter dem Agenten

Sind bezüglich eines neuen Objekts deiktische räumliche Relationen zu bestimmen, so ist das Referenzobjekt der Agent. Deiktische räumliche Relationen repräsentieren die Lage eines Objektes relativ zur Hauptausrichtung des Agenten, weswegen sie bezüglich eines primären oder sekundären Referenzsystems evaluiert werden.<sup>53</sup> Die Eingabe der entsprechenden Funktion zur Bestimmung der bestanzwendbaren deiktischen Relation ist die Funktion *Evaluate-Deictic-Rels*

<sup>52</sup>Eine dedizierte Analyse der formalen Berechnung topologischer Relationen liegt außerhalb des Kontext dieser Arbeit. Verschiedene Berechnungsmodelle topologischer Relationen sind in Kapitel 2.3.2 referenziert.

<sup>53</sup>Zur Evaluation von räumlichen Relationen basierend auf geometrischen Objektrepräsentationen (vgl. [Her96]).

(vgl. Funktion 24).

(24) Evaluate-Deictic-Rels(lo, ego)

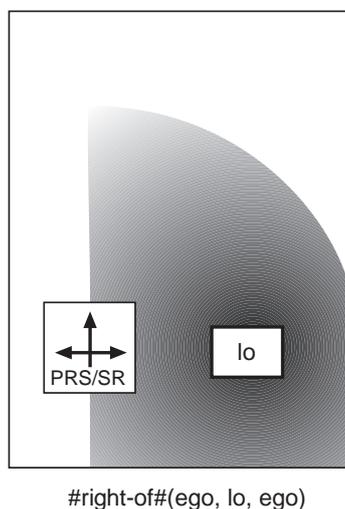


Abbildung 4.45: Entsprechend der Funktion  $\#right-of\#(ego, lo, ego)$  konstituiertes Gebiet der Anwendungsgrade

Einer Illustration des Zusammenhangs zwischen Objekten, Referenzsystemen und Anwendungsgraden im deiktischen Falle, dient die Abbildung 4.45. Das Gebiet der Anwendungsgrade ist durch den schattierten Bereich um die Lokation des Betrachters indiziert. Dunkle Gebiete indizieren hohe, helle geringere Anwendungsgrade. Im Falle der deiktischen Evaluation ist das Anwendungsgebiet im PRS bzw. im SRS verankert. Im vorliegenden Fall besteht eine hohe Anwendbarkeit der Relation  $\#right-of\#$  für das Objekt  $lo$ .

**4.2.4.2.4 Intrinsische räumliche Relationen** Die Evaluation intrinsischer räumlicher Relationen setzt die Etablierung eines virtuellen Referenzsystems voraus. Die Menge der hier betrachteten intrinsischen Relationen stimmt mit den deiktischen Relationen überein, mit dem Unterschied, daß das Referenzsystem sich an einer externen Lokation befindet:<sup>54</sup>

1.  $\#left-of\#(ego, lo, refo)$ :  $lo$  befindet sich links vom Referenzobjekt
2.  $\#right-of\#(ego, lo, refo)$ :  $lo$  befindet sich rechts vom Referenzobjekt
3.  $\#behind\#(ego, lo, refo)$ :  $lo$  befindet sich hinter dem Referenzobjekt
4.  $\#in-front-of\#(ego, lo, refo)$ :  $lo$  befindet sich vor dem Referenzobjekt
5.  $\#above\#(ego, lo, refo)$ :  $lo$  befindet sich über dem Referenzobjekt
6.  $\#under\#(ego, lo, refo)$ :  $lo$  befindet sich unter dem Referenzobjekt

<sup>54</sup>Der hier verwendete Ansatz erlaubt es, nur solche intrinsische Relationen zu berechnen, die bezüglich einem *im* Referenzobjekt verankerten virtuellen Referenzsystem definiert sind (vgl. [Gap94]).

Die Funktion zur Evaluation der bestmöglichen intrinsischen Relation zwischen der Lokation des Betrachters, dem Lokalisations- und dem Referenzobjekt ist *Evaluate-Intrinsic-Rel*s (vgl. Funktion 25). Ebenso wie bei deiktischen Relationen, liefert die Evaluation einer intrinsischen Relation als Rückgabewert einen Anwendbarkeitsgrad.

(25) Evaluate-Intrinsic-Rel(s(ego, lo, refo))

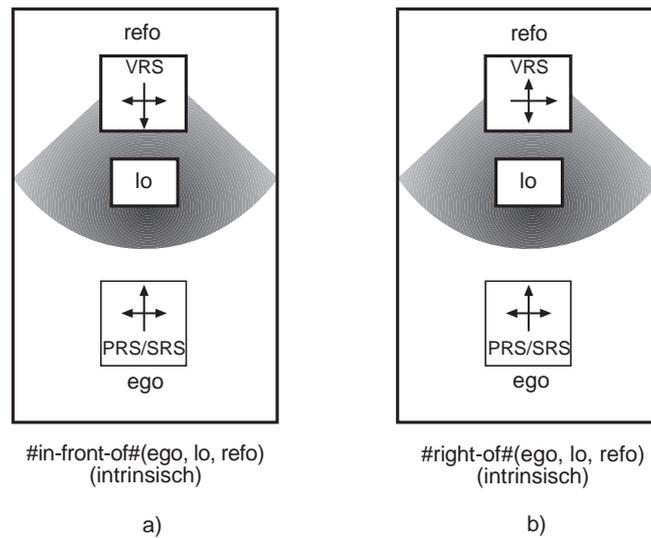


Abbildung 4.46: Evaluation höchstanwendbarer intrinsischer Relationen: a) bei antiparallelem VRS, b) bei orthogonal-orientiertem VRS

Durch die einzelsprachabhängige Festlegung des virtuellen Referenzsystems, ergibt die Evaluation intrinsischer Relationen unterschiedliche Relationen mit höchstem Anwendbarkeitsgrad (vgl. Abbildung 4.46).<sup>55</sup> Im ersten Fall (vgl. Abbildung 4.46a)) ist das virtuelle Referenzsystem antiparallel zum primären bzw. sekundären Referenzsystem des Betrachters. Die höchstanwendbare intrinsische Relation ist dabei #in-front-of#. <sup>56</sup> Im zweiten Fall (Abbildung 4.46b) ist das VRS nach rechts orientiert, so daß in dieser Situation die Relation #right-of# den höchsten Anwendbarkeitsgrad besitzt.

Zusammenfassend dienen die hier betrachteten räumlichen Relationen der Integration von einzelnen Objekten zu einer propositionalen Repräsentation eines räumlichen Sachverhalts. Topologische Relationen werden dazu verwendet, um physikalische Kontaktbeziehungen zu repräsentieren, wobei im weiteren alleinig die Relation #meet# verwendet wird. Deiktische und intrinsische Relationen basieren im Gegensatz zu topologischen Relationen auf der Verwendung eines Referenzsystems. Die räumlichen Beziehungen zwischen einem einzelnen Objekt und dem

<sup>55</sup>In diesem Beispiel wird zur besseren Illustration davon ausgegangen, daß das virtuelle Referenzsystem durch intrinsische Eigenschaften des Referenzobjektes in eine antiparallele bzw. orthogonale Orientierung relativ zum PRS/SRS rotiert wird. Eine realweltliche Instantiierung dieses Beispiels ist gegeben, wenn das Referenzobjekt eine Kirche mit Portal an der Seite zum Lokalisationsobjekt ist (vgl. Abbildung 4.46a), bzw. wenn sich das Portal an der rechten Seite befindet (vgl. Abbildung 4.46b).

<sup>56</sup>Dieser Fall, wenn VRS und PRS/SRS antiparallel sind, ist identisch mit einer deiktischen Evaluation. Eine effizientere Evaluation deiktischer und intrinsischer Relationen wäre in diesem Falle über die Auswertung räumlicher linearer Ordnungen und einer Evaluation einer Relation der *räumlichen Nähe* denkbar, die über die Distanz vom Betrachter definiert ist. Sind VRS und PRS/SRS parallel, so kehrt sich die Reihenfolge in der linearen Ordnung nur um. Dadurch lassen sich der deiktische und der intrinsische Fall zusammenfassen.

Betrachter werden auf der Basis eines egozentrischen Referenzsystems durch deiktische Relationen repräsentiert. Hingegen werden intrinsische Relationen bezüglich eines allozentrischen, virtuellen Referenzsystems, verwendet, um räumliche Lagebeziehungen zwischen zwei Objekten zu repräsentieren.

Sowohl für deiktische, als auch für intrinsische Relationen wird von einer minimalen Menge von Relationen ausgegangen, die bezüglich kartesisch definierter Referenzsysteme abgeleitet werden. Komplexere Relationen können aus diesen Minimalmengen abgeleitet werden. Mit den verschiedenen Referenzsystemen und den einzelnen Relationentypen lassen sich integrierte räumliche Repräsentationen komplexer Verkehrssituationen ableiten. Im folgenden Kapitel wird eingehend vorgestellt, welche räumlichen Repräsentationen aufgebaut werden, die letztendlich eine sprachliche Beschreibung eines Wegabschnitts ermöglichen.

#### 4.2.4.3 Der räumliche Kurzzeitspeicher

In den vorangegangenen Abschnitten ist vorgestellt worden, wie Pfadelemente gesucht (vgl. Kapitel 4.2.2), visuell saliente Objekte ausgewählt (vgl. Kapitel 4.2.3), sowie Referenzsysteme (vgl. Kapitel 4.2.4.1) und räumliche Relationen (vgl. Kapitel 4.2.4.2) etabliert werden. Im weiteren wird diskutiert, wie alle diese Einzelkonzepte integriert werden, um eine räumliche Vorstellung einer Situation, eine sogenannte *räumliche Konfigurationsbeschreibung*, zu etablieren. Bedingt durch die Dynamik einer Situation, wird eine solche räumliche Konfigurationsbeschreibung inkrementell evaluiert und angepaßt. Das dadurch etablierte raumbezogene Wissen über eine Situation wird im sogenannten *räumlichen Kurzzeitspeicher* verwaltet.

Im Vordergrund dieser Arbeit steht die Bestimmung einer *minimalen* Menge an räumlichem Wissen, das für die Generierung inkrementeller Wegbeschreibungen notwendig ist. Problematisch dabei ist es vor allem, aus der Menge der potentiell etablierbaren räumlichen Relationen geeignete auszuwählen. Hierzu wird ein inkrementeller Ansatz verwendet, der, ausgehend von minimalem räumlichen Wissen, dieses schrittweise erweitert. Dies gewährleistet einerseits eine effiziente Bereitstellung solchen Wissens und andererseits die flexible Adaptation an ressourcenbedingte Beschränkungen. Operational gesehen, wird in einem ersten Schritt eine *minimale Konfigurationsbeschreibung* aufgebaut. Diese setzt den Agenten durch Etablierung von deiktischen Relationen in eine räumliche Lagebeziehung zum aktuellen Pfadabschnitt, auf dem er sich gerade befindet, sowie zu dem nächsten Verzweigungspunkt und dem anschließenden Pfadabschnitt. Hiernach werden zwischen den einzelnen Pfadelementen intrinsische Relationen bestimmt, die dazu verwendet werden, *pfadbezogene Intentionen* abzuleiten. Eine minimale räumliche Konfigurationsbeschreibung ist ausreichend, um eine minimale Beschreibung eines Weges zu generieren. Ist der Zeitrahmen ausreichend groß und gibt es in der Situation saliente Landmarken, wird die räumliche Konfigurationsbeschreibung durch Integration dieser Landmarken erweitert. Zwischen Objekten der Konfigurationsbeschreibung und der zu integrierenden Landmarke werden intrinsische Relationen mustergesteuert etabliert. Eine dadurch entstehende *erweiterte räumliche Konfigurationsbeschreibung* ist Ausgangspunkt komplexerer Wegbeschreibungen.

Im weiteren Verlauf werden räumliche Konfigurationsbeschreibungen genauer diskutiert. Zuerst wird gezeigt, wie minimales, räumliches Routenwissen integriert wird. Anschließend wird vorgestellt, wie diese Repräsentation einer Situation durch Landmarkenwissen erweitert werden kann.

**4.2.4.3.1 Räumliche Konfigurationsbeschreibungen** Um in dynamischen Situationen Raumbeschreibungen generieren zu können, muß der Agent über lokal-stabiles, raumbezogenes Wissen einer Situation verfügen. Entsprechend dem Ansatz, der zur Auswahl salienter Objekte vorgestellt worden ist (vgl. Kapitel 4.2.3), wird raumbezogenes Wissen bezüglich diskreten

Zeitpunkten etabliert. Das Modell basiert auf einem zweistufigen Ansatz (vgl. Abbildung 4.47). Während der ersten Phase wird Routenwissen durch Integration von Pfadelementen einer Situation aufgebaut, wodurch eine minimale räumliche Konfigurationsbeschreibung etabliert wird. Falls Landmarken ausgewählt worden sind, werden diese im zweiten Schritt zu Entitäten der räumlichen Konfigurationsbeschreibung in Beziehung gesetzt. Bezüglich einer Situation wird die zweite Phase solange durchlaufen, bis entweder keine weiteren salienten Landmarken mehr vorhanden sind oder Ressourcenbeschränkungen dies verhindern.

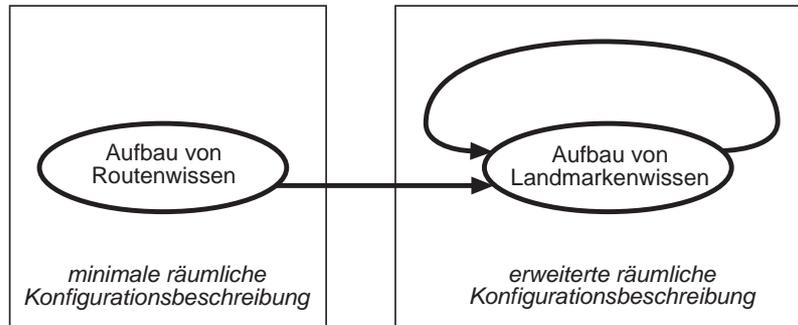


Abbildung 4.47: Zweistufige Generierung raumbezogenen Wissens in dynamischen Situationen

Eine *räumliche Konfigurationsbeschreibung* (RKB) repräsentiert auf qualitative Weise visuell-räumliche Sachverhalte einer Situation. Gebietskonstituierende, räumliche Relationen, so wie sie in Kapitel 4.2.4.2 vorgestellt worden sind, sind weitgehend invariant bezüglich geringer räumlicher Lageänderungen, wodurch sich eine Stabilität einer RKB ergibt. Bewegt sich der Agent translatorisch, so selektiert er neue Entitäten aus der sich ändernden Situation. Entitäten, die nicht mehr sichtbar sind, werden zusammen mit den räumlichen Relationen aus der RKB entfernt (vgl. Kapitel 3).

Sobald der Agent eine neue Situation erreicht, wird die RKB initialisiert.<sup>57</sup> Dies ist eine Einschränkung der Allgemeingültigkeit bezüglich der Verarbeitung räumlichen Wissens. Hingegen ist dies in der Domäne von inkrementellen Wegbeschreibungen legitim, da in keiner der Beschreibungen aus den Untersuchungen Bezüge zu räumlichen Wissen aus vorangegangenen Pfadabschnitten gemacht worden sind.

**4.2.4.3.1.1 Minimale räumliche Konfigurationsbeschreibungen** Jeder Verzweigungspunkt beendet den aktuell verfolgten Pfadabschnitt und veranlaßt den Agenten, sich und dadurch auch das primäre Referenzsystem in eine bestimmte Richtung zu drehen, um dann auf diesem Pfadabschnitt die Navigation fortzusetzen. Zwischen zwei Pfadabschnitten befindet sich grundsätzlich ein Verzweigungspunkt (vgl. Kapitel 4.2.2). Eine physische Grundbedingung an Pfadelemente<sup>58</sup> ist, daß der Übergang zwischen einem Pfadabschnitt und einem Verzweigungspunkt stetig<sup>59</sup> ist und eine Kontaktrelation erfüllt ist, die durch die topologische Relation *#meet#* repräsentiert wird. Die Integration eines Pfadelementes *p* in eine räumliche Konfi-

<sup>57</sup>Eine natürliche Fortsetzung dieser Arbeit ist es, das Wissen einer räumlichen Konfigurationsbeschreibung durch geeignete Funktionen an eine Langzeitspeicherung in Form einer kognitiven Karte weiterzuleiten (vgl. Kapitel 2.3.7, 2.6).

<sup>58</sup>Die Menge der Pfadelemente ist die Gesamtheit aus Pfadabschnitten und Verzweigungspunkten (vgl. Kapitel 4.2.2).

<sup>59</sup>Der Aspekt der Stetigkeit wird im weiteren Verlauf nicht weiter betrachtet. Ein geeignetes Konzept der *qualitativen Stetigkeit* ist eng mit dem Energieaufwand verknüpft, den der Agent benötigt, um den Übergang erfolgreich zu überwinden. Sprünge bei Übergängen, wie bei Treppen, müssen bis zu einem gewissen Grade als qualitativ stetig konzeptualisiert werden.



befindet (vgl. Relationen in 27).

$$(27) \quad \#left-of\#(ego, S2, ego) \ \& \ \#in-front-of\#(ego, S2, ego)$$

Zum Verzweigungspunkt  $C$  läßt sich wiederum nur eine deiktische Relation etablieren, die repräsentiert, daß  $C$  sich „vor“ dem Agenten befindet (vgl. Relationen in 28).

$$(28) \quad \#in-front-of\#(ego, C, ego)$$

Zur Evaluierung physikalischer Begebenheiten wird zwischen dem Agenten und dem Pfadabschnitt  $S1$  evaluiert, ob eine Kontaktrelation besteht (vgl. Abbildung 4.50a) und die Relationen in 29).

$$(29) \quad \#meet\#(ego, S1)$$

Weiterhin bestehen Kontaktrelationen zwischen den Pfadelementen (vgl. Relationen in 30).

$$(30) \quad \#meet\#(S2, C), \ \#meet\#(C, S1)$$

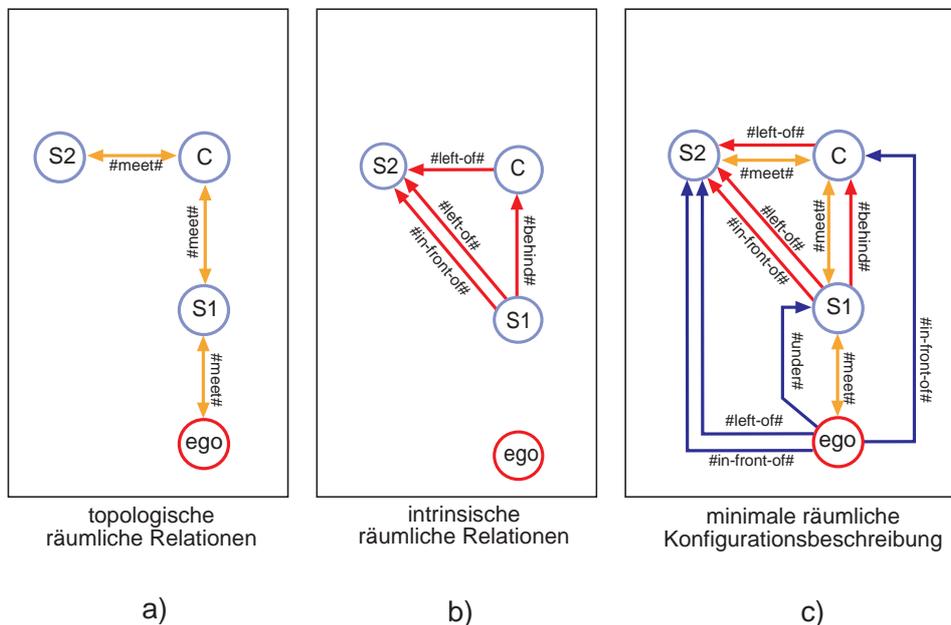


Abbildung 4.50: a) Topologische Relationen, b) intrinsic Relationen zwischen Pfadelementen, c) minimale räumliche Konfigurationsbeschreibung

Das Wissen, was alleine durch deiktische und topologische Relationen repräsentiert wird, ist als Grundlage für sprachliche Beschreibungen nicht ausreichend, da u. a. die *Transitivitätsei-*

*genschaft* zwischen Pfadabschnitten und Verzweigungspunkten nicht gewährleistet ist.<sup>60</sup> Eine Situation besteht grundsätzlich aus mindestens zwei Pfadabschnitten und einem dazwischenliegenden Verzweigungspunkt. Es werden zwischen zwei Pfadelementen einer Situation dann intrinsische Relationen etabliert, wenn sie durch eine Kontaktrelation miteinander verbunden sind, oder wenn es sich um zwei Pfadabschnitte handelt. In Abbildung 4.49 liegt *S2* sowohl „links“ und „hinter“ *S1* (vgl. Abbildung 4.50b) und die Relationen in 31).

$$(31) \quad \#left-of\#(ego, S2, S1) \ \& \ \#behind\#(ego, S2, S1)$$

Ferner gilt, daß aus Sicht des Agenten die Kreuzung *C* „hinter“ *S1* und das *S2* „links von“ *C* liegt (vgl. Relationen in 32).

$$(32) \quad \begin{array}{l} \#behind\#(ego, C, S1) \\ \#left-of\#(ego, S2, C) \end{array}$$

Insgesamt besteht eine minimale räumliche Konfigurationsbeschreibung aus der Gesamtheit der räumlichen Relationen mit den höchsten Anwendbarkeitsgraden, so wie es in Abbildung 4.50c) indiziert ist.

**4.2.4.3.1.2 Pfadgeleitete Intentionenmuster** Aus den intrinsischen Relationen zwischen Pfadelementen lassen sich pfadgeleitete Intentionen des Agenten bezüglich einer Situation *S* ableiten. Durch die qualitative räumliche Repräsentation der Pfadelemente zueinander lassen sich *Intentionenmuster* (*IM*) angeben, nach denen entschieden werden kann, in welche Richtung der Pfad weiter verfolgt wird. Basierend auf den intrinsischen Relationen zwischen Pfadelementen konstituieren sich sechs Intentionenmuster, denen *pfadgeleitete Intentionen* (*PI*) zugeordnet sind (vgl. Abbildung 4.51). Entsprechend den kartesischen Dimensionen der Referenzsysteme werden folgende sechs<sup>61</sup> pfadgeleitete Intentionen unterschieden:

1. *nach-rechts-bewegen*: der nachfolgende Pfadabschnitt *S2* liegt qualitativ „rechts“ vom Verzweigungspunkt *C* und dem aktuellen Pfadabschnitt *S1* ( $\$turn-right\$(S1, C, S2)$ ).<sup>62</sup>
2. *nach-links-bewegen*: *S2* liegt qualitativ „links“ von *C* und *S1* ( $\$turn-left\$(S1, C, S2)$ ).
3. *geradeaus-bewegen*: *S2* liegt qualitativ „hinter“ *C* und *S1* ( $\$go-straight\$(S1, C, S2)$ ).
4. *nach-oben-bewegen*: *S2* liegt qualitativ „über“ *C* und *S1* ( $\$go-up\$(S1, C, S2)$ ).
5. *nach-unten-bewegen*: *S2* liegt qualitativ „unter“ *C* und *S1* ( $\$go-down\$(S1, C, S2)$ ).
6. *horizontal-bewegen*: *S2* liegt qualitativ „eben“ bezüglich *C* und *S1* ( $\$go-plane\$(S1, C, S2)$ ).

Da intrinsische Relationen zwischen Pfadelementen relativ zum primären Referenzsystem etabliert sind, folgt eine ebensolche Verankerung pfadgeleiteter Intentionen. Die Bestimmung einer *PI* beruht auf einer schwellwertbasierten Auswertung des assoziierten Intentionenmusters (vgl.

<sup>60</sup>Liegt ein Pfadelement *A* links sowie die Elemente *B* und *C* vor dem Agenten, so kann nicht eindeutig abgeleitet werden, wie *A* bezüglich *B* liegt.

<sup>61</sup>Die Möglichkeit einer Rückwärtsbewegung wird nicht betrachtet.

<sup>62</sup>Pfadgeleitete Intentionenmuster sind durch \$ prä- und postfigiert.



(vgl. Abbildung 4.51d)-f) und Gleichung 34) bezogen ist, zuordnen.

$$(34) \quad \begin{aligned} PI_{\text{horizontal}} &= \{\text{\$turn-left\$}, \text{\$turn-right\$}, \text{\$go-straight\$}\} \\ PI_{\text{vertikal}} &= \{\text{\$go-up\$}, \text{\$go-down\$}, \text{\$go-plane\$}\} \end{aligned}$$

Zwei pfadgeleitete Intentionen aus verschiedenen Klassen sind *unabhängig* voneinander, wohingegen solche aus einer Klasse dies nicht sind. Zwei unabhängige pfadgeleitete Intentionen lassen sich durch Konjunktionsbildung zu *konjungierten pfadgeleiteten Intentionen* (KPI) integrieren. Dadurch ist es beispielsweise möglich, eine Intention nach „links hoch“ zu repräsentieren. Andererseits werden Intentionen wie eine intendierte Bewegung nach „links rechts“ ausgeschlossen.

Konjungierte PI	Konjunktion linear unabhängiger PIs
1. $\text{\$turn-left\$}$	$\text{\$turn-left\$} \wedge \text{\$go-plane\$}$
2. $\text{\$turn-left-up\$}$	$\text{\$turn-left\$} \wedge \text{\$go-up\$}$
3. $\text{\$turn-left-down\$}$	$\text{\$turn-left\$} \wedge \text{\$go-down\$}$
4. $\text{\$turn-right\$}$	$\text{\$turn-right\$} \wedge \text{\$go-plane\$}$
5. $\text{\$turn-right-up\$}$	$\text{\$turn-right\$} \wedge \text{\$go-up\$}$
6. $\text{\$turn-right-down\$}$	$\text{\$turn-right\$} \wedge \text{\$go-down\$}$
7. $\text{\$go-straight\$}$	$\text{\$go-straight\$} \wedge \text{\$go-plane\$}$
8. $\text{\$go-up\$}$	$\text{\$go-straight\$} \wedge \text{\$go-up\$}$
9. $\text{\$go-down\$}$	$\text{\$go-straight\$} \wedge \text{\$go-down\$}$

Tabelle 4.2: Konjunktionen von linear unabhängigen pfadgeleiteten Intentionen zu konjungierten pfadgeleiteten Intentionen

Durch Konjunktion von pfadgeleiteten Intentionen aus  $PI_{\text{horizontal}}$  und  $PI_{\text{vertikal}}$  lassen sich neun konjungierte pfadgeleitete Intentionen ableiten (vgl. Tabelle 4.2). Wie aus Tabelle 4.2 ersichtlich ist, besitzen die Klassen  $PI_{\text{horizontal}}$  mit  $\text{\$go-straight\$}$  und  $PI_{\text{vertikal}}$  mit  $\text{\$go-plane\$}$  neutrale Elemente. Werden diese beiden  $PI$ s miteinander konjungiert (Zeile 7 in Tabelle 4.2), so dominiert die Intention aus  $PI_{\text{horizontal}}$ . Per Durchschnittsbildung des Anwendbarkeitsgrades von  $PI(AG_{\text{int}}(IM_1, S))$  und  $PI(AG_{\text{int}}(IM_2, S))$ , ergibt sich der Anwendbarkeitsgrad der *konjungierten pfadgeleiteten Intention* (KPI)  $AG_{\text{kpi}}(IM_1, IM_2, S)$  (vgl. Gleichung 35).

$$(35) \quad AG_{\text{kpi}}(IM_1, IM_2, S) = \frac{1}{2}(AG_{\text{int}}(IM_1, S) + AG_{\text{int}}(IM_2, S))$$

Zusammenfassend heißt dies, daß durch Etablierung von räumlichen Relationen zwischen den aktuellen Pfadelementen und der Lokation des Agenten eine minimale räumliche Konfigurationsbeschreibung aufgebaut wird. Von dieser Repräsentation leitet sich die aktuelle pfadgeleitete Intention des Agenten ab. Wie Landmarken in räumliche Konfigurationen integriert werden, wird im folgenden vorgestellt.

**4.2.4.3.1.3 Erweiterte räumliche Konfigurationsbeschreibungen** In der zweiten Phase des Aufbaus raumbezogenen Wissens (vgl. Abbildung 4.47) werden Landmarken integriert, die aufgrund ausreichender visueller Salienz ausgewählt worden sind (vgl. Kapitel 4.2.3). Landmarken dienen als räumliche Referenzpunkte in einer Situation und sind von pfadgeleiteten

Intentionen abhängig.

Wie aus den Beschreibungen des Korpus hervorgeht (vgl. Kapitel 3), wird bei einer pfadbezogenen Intention des Abbiegens<sup>63</sup> ein Bereich des Verzweigungspunktes fokussiert, der in der Nähe des nachfolgenden Pfadabschnitts liegt. Findet das Objektauswahlmodul (vgl. Kapitel 4.2.3) in diesem räumlichen Aufmerksamkeitsbereich (RAB) eine Landmarke, so wird bezüglich dieser ein sekundäres Referenzsystem etabliert, falls der Rotationswinkel  $\alpha$  im Elastizitätsbereich liegt (vgl. Abbildung 4.53).

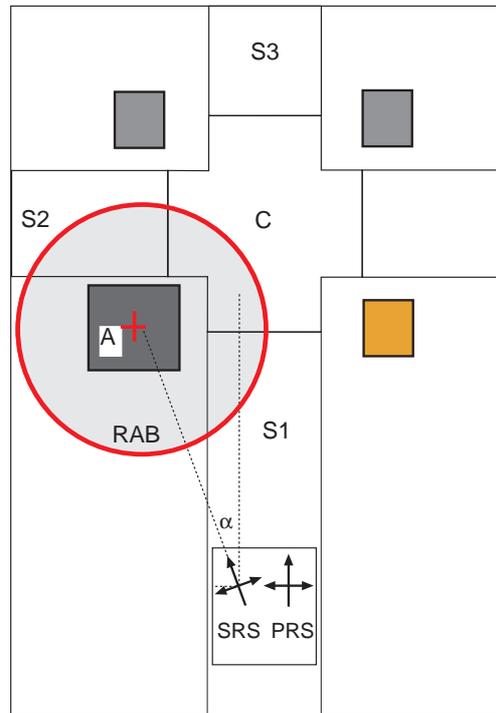


Abbildung 4.53: Konstruktion eines sekundären Referenzsystems in Abhängigkeit von einer pfadgeleiteten Intention des Linksabbiegens

Ist eine Landmarke in einer Situation ausgewählt worden, so wird sie in die räumliche Konfigurationsbeschreibung integriert. Ebenso wie dies im ersten Schritt des Aufbaus raumbezogenen Wissens der Fall ist, werden bezüglich dieser Landmarke zuerst deiktische Relationen etabliert (vgl. Abbildung 4.54). Kontaktrelationen sind nicht zu betrachten, da der Agent mit der Landmarke selbst nicht in Interaktion tritt. Im darauffolgenden Schritt wird der *räumliche Anwendbarkeitsgrad* einer Landmarke bezüglich der Situation bestimmt, der angibt, inwieweit die Landmarke geeignet ist, um als Referenzobjekt verwendet zu werden. Zuletzt werden intrinsische Relationen zum aktuellen und folgenden Pfadabschnitt sowie zur Kreuzung etabliert. Eine funktionale Beschreibung findet sich in Funktion 10.

Die Evaluation von deiktischen Relationen einer Landmarke  $A$  erfolgt bezüglich eines geeigneten sekundären Referenzsystems  $SRS$  (vgl. Abbildung 4.53). In diesem Beispiel werden zwischen der Landmarke und der Lokation des Agenten zwei deiktische Relationen etabliert, welche repräsentieren, daß sich  $A$  „links vor“ dem Agenten befindet (vgl. Relationen 36).

$$(36) \quad \#left-of\#(ego, A, ego) \ \& \ \#in-front-of\#(ego, A, ego)$$

<sup>63</sup>Solche  $PIs$  enthalten in ihrer Definition  $\$turn-left\$$  oder  $\$turn-right\$$ .

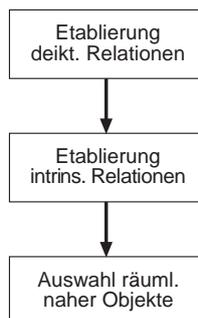


Abbildung 4.54: Stufen zur Integration einer Landmarke

## Funktion 10: Integrate-Landmarks

```

Integrate-Landmarks(S, PRS, focus, lm)
  Evaluate-Deictic-Rels(SRS, lm);
  vrs := Select-Virtual-Frame-of-Reference(srs, lm);
  Determine-Spatially-Near-Objects(lm);
  for all o in RAGM(lm) do
    Evaluate-Intrinsic-Rels(vrs, lm, o);
  Det-Appl-of-LM-in-RKB(lm)
  
```

Eine intrinsische Eigenschaft der zu integrierenden Landmarke ist ihre visuelle Salienz, welche sie zur Verwendung als Referenzobjekt auszeichnet. Bei der Evaluation von intrinsischen räumlichen Relationen wird die neue Landmarke als Referenzobjekt verwendet. Aus dem gegebenen Blickwinkel des Agenten wird zur Evaluation geeigneter intrinsischer räumlicher Relationen ein geeignetes virtuelles Referenzsystem bestimmt. Visuelle Salienz ist keine *richtungsgebende* intrinsische Eigenschaft einer Landmarke. Besitzt eine Landmarke keine richtungsgebende intrinsische Eigenschaft, so wird, entsprechend den in Kapitel 4.2.4.1 vorgestellten Regeln, ein virtuelles Referenzsystem (VRS) etabliert.

Die bezüglich Abbildung 4.53 evaluierten räumlichen Relationen sind in Abbildung 4.55 eingezeichnet.<sup>64</sup> Danach befindet sich der Pfadabschnitt *S1* rechts von *A*, der Verzweigungspunkt *C* rechts hinter *A* und der Pfadabschnitt *S2* links hinter *A* (vgl. die intrinsischen Relationen 37).

$$(37) \quad \begin{array}{l} \#right-of\#(ego, A, S1) \\ \#right-of\#(ego, A, C) \ \& \ \#behind\#(ego, A, C) \end{array}$$

Ob eine Landmarke als Referenzobjekt zur Beschreibung einer Aktion geeignet ist, hängt nicht nur von visuellen Eigenschaften ab. Vielmehr ist zusätzlich notwendig, daß sie sich an einer geeigneten räumlichen Lokation befindet. Wie aus den Experimenten hervorgeht (vgl. Kapitel 3), werden Landmarken nur dann in die Beschreibung integriert, wenn sie sich in der Halbebene befinden, die durch das nächste Pfadelement determiniert ist. Zusätzlich müssen sie sich in der Nähe der Pfadelemente befinden.

<sup>64</sup>Es werden der Klarheit wegen für Landmarken extrinsisch determinierte, virtuelle Referenzsysteme angenommen.



Die Berechnung der Anwendbarkeit einer Landmarke erfolgt musterbasiert. Aufbauend auf den möglichen Lokationen potentieller Referenzobjekte, lassen sich für jede Abbiegeintention die obengenannten zwei Fälle unterscheiden, die verschiedene *räumliche Anwendbarkeitsgradmuster* (RAGM) besitzen (vgl. Tabelle 4.3). Landmarken, die sich vor S2 befinden, sind durch  $LM_{\text{front}}$  und solche hinter S2 durch  $LM_{\text{back}}$  indiziert.

RAGM(\$\text{turn-left}\$)	S1	C	S2
$LM_{\text{front}}$	#right-of#	#right-of# #behind#	#behind#
$LM_{\text{back}}$	#right-of#	#in-front-of#	#in-front-of#

RAGM(\$\text{turn-right}\$)	S1	C	S2
$LM_{\text{front}}$	#left-of#	#left-of# #behind#	#behind#
$LM_{\text{back}}$	#left-of#	#in-front-of#	#in-front-of#

Tabelle 4.3: Musterbasierte Auswahl von potentiellen Referenzobjekten

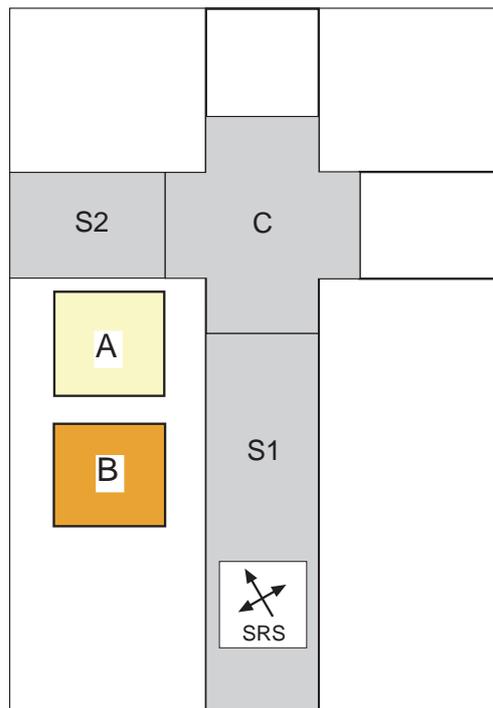


Abbildung 4.57: Zwei potentielle Referenzobjekte A und B

Die Evaluation der Anwendbarkeit einer Landmarke erfolgt über die räumlichen Anwendbarkeitsgrade der einzelnen räumlichen Relationen. Der *räumliche Anwendbarkeitsgrad*  $RAG(lm)$  einer Landmarke ergibt sich aus dem arithmetischen Mittel der Quadrate räumlicher Anwendbarkeitsgrade der räumlichen Relationen, die durch das betreffende räumliche Anwendbarkeitsgrad-

muster vorgegeben sind (vgl. Gleichung 38).<sup>65</sup> Ein geeignetes  $n$  ergibt sich aus der anwendbaren Eintrag aus den Tabellen 4.3. Befindet sich die Landmarke vor dem nächsten Pfadabschnitt, so ist  $n$  gleich vier, liegt sie dahinter, so ist  $n$  gleich 3 (vgl. Tabelle 4.3).

$$(38) \quad \text{RAG}(\text{lm}) = \frac{1}{n} \sum_{i=1}^n ap - \text{rel}(R_i(\text{ego}, \text{lm}, \text{obj}_i))^2$$

In Abbildung 4.57 erfüllen beide Landmarken  $A$  und  $B$  die Bedingung in der geeigneten Halbebene zu liegen. Beide Landmarken liegen vor dem nächsten Pfadabschnitt  $S2$ , so daß das erste räumliche Anwendbarkeitsgradmuster von  $\text{RAGM}(\text{\$turn-left}\$)$  anwendbar ist. Wie sich aus den folgenden Berechnungen ergibt, besitzt die Landmarke  $A$  einen höheren räumlichen Anwendbarkeitsgrad  $\text{RAG}$  als  $B$ .

$$(39) \quad \begin{aligned} \text{RAG}(A) &= (\#\text{right-of}\#(\text{ego}, A, S1)^2 + \#\text{right-of}\#(\text{ego}, A, C)^2 + \\ &\quad \#\text{behind}\#(\text{ego}, A, C)^2 + \#\text{behind}\#(\text{ego}, A, S2)^2)/4 \\ &= (1^2 + 0, 7^2 + 0, 7^2 + 0, 9^2)/4 = 0, 6975 \end{aligned}$$

$$(40) \quad \begin{aligned} \text{RAG}(B) &= (\#\text{right-of}\#(\text{ego}, B, S1)^2 + \#\text{right-of}\#(\text{ego}, B, C)^2 + \\ &\quad \#\text{behind}\#(\text{ego}, B, C)^2 + \#\text{behind}\#(\text{ego}, B, S2)^2)/4 \\ &= (0, 9^2 + 0, 5^2 + 0, 4^2 + 1^2)/4 = 0, 555 \end{aligned}$$

Hieraus folgt, daß Landmarke  $A$  bezogen auf die räumliche Lokation als Referenzobjekt zu präferieren ist. Wie aber später bei der vollständigen Beschreibung der Auswahl von Referenzobjekten gezeigt wird, sind die visuelle Salienz und der räumliche Anwendbarkeitsgrad einer Landmarke zu integrieren (vgl. Kapitel 4.2.6).

In diesem Kapitel sind die Konzepte vorgestellt worden, mit denen der Agent räumliches Wissen über eine Situation repräsentiert. Bedingt durch Beschränkungen, die durch die Ressource *Zeit* und durch Speicherplatzrestriktionen vorgegeben sind, wird nur eine minimale Menge an räumlichen Relationen etabliert, die zur Handlungsfindung notwendig sind. Bezüglich einer Situation wird zuerst eine minimale räumliche Konfigurationsbeschreibung generiert, die die relevanten Straßenelemente zum Agenten und zueinander in räumliche Beziehung setzt. Eine minimale räumliche Konfigurationsbeschreibung enthält das minimale räumliche Wissen, welches notwendig ist, um einen Wegabschnitt adäquat zu beschreiben. Steht dem Agenten genügend Zeit zur Verfügung, kann die räumliche Konfigurationsbeschreibung durch Landmarkenwissen erweitert werden. Eine solche Repräsentation wird als erweiterte räumliche Konfigurationsbeschreibung bezeichnet.

Bevor die Verbindung zwischen räumlichem Wissen und sprachlichen Beschreibungen vorgestellt wird, diskutiert das folgende Kapitel die qualitativen Navigationsoperationen, die dem Agenten zur Eigenbewegung zur Verfügung stehen.

---

<sup>65</sup>Bezogen auf den vorliegenden Korpus hat sich diese Definition des Anwendbarkeitsgrades einer Landmarke als geeignet herausgestellt. Dies liegt u.a. an der Eigenschaft der Quadratfunktion, daß der relative Unterschied zwischen kleineren und größeren Werten vergrößert wird. Ferner besitzt die Funktion die gewünschten Eigenschaften streng monoton steigend und positiv zu sein. Weitere, dediziertere empirische Untersuchungen müssen zeigen, ob dieser Zusammenhang allgemein gilt.

### 4.2.5 Funktionen des Navigationsmoduls

Durch das Navigationsmodul werden qualitative Navigationsoperationen evaluiert. Sie erlauben es dem Agenten, sich relativ zu seiner Umgebung fortzubewegen. Als Rückgabe liefert die Ausführung einer Navigationsfunktion an das Raumintegrationsmodul die Lokation und die Orientierung relativ zur Umgebung zurück. Diese Information erneuert die Verankerung und Orientierung des primären Referenzsystems (vgl. Abbildung 4.58). Zwei navigationsorientierte Funktionen stehen dem Agenten zur Verfügung, die es ihm erlauben, eine translatorische (MOVE) oder rotationale Bewegung (ROTATE) auszuführen.

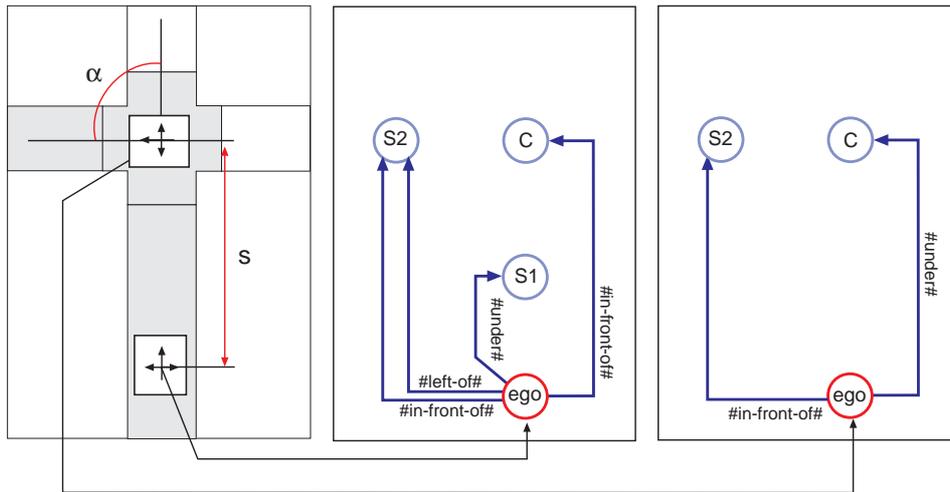


Abbildung 4.58: Ableitung der Lokation und Orientierung des *PRS* aus der Information des motorischen Moduls bezüglich Lokationen, die sich in einer Distanz  $s$  vor der Kreuzung und auf der Kreuzung befinden.

In der betrachteten stadtähnlichen Domäne können Bewegungen ausschließlich auf Pfadelementen ausgeführt werden. Sind der aktuelle, der nachfolgende Pfadabschnitt sowie der Verzweigungspunkt festgelegt und pfadbasierte Intentionen abgeleitet worden, so erhält das Navigationsmodul vom *RI-M* die Anweisung, entsprechend der topographischen Lagebeziehungen der Pfadelemente geeignete Navigationsfunktionen zu evaluieren. Eine translatorische Bewegung wird durch Angabe einer Geschwindigkeit  $v$  und einer Länge der zu bewältigenden Distanz  $s$  aktiviert ( $MOVE(v, s)$ ). Die Aktivierung einer Rotation des gesamten Agenten<sup>66</sup> erfolgt über die Angabe der Rotationsgeschwindigkeit  $r$  und dem Rotationswinkel  $\alpha$  ( $ROTATE(r, \alpha)$ ).

Während der Evaluation einer Bewegungsfunktion erhält das *RI-M* in diskreten Abständen die Orientierung und die Lokation des Agenten (vgl. Abbildung 4.58). In den Abbildungen 4.59 und 4.60 sind Beispiele für die Sequenz einer translatorischen und einer rotationalen Aktion illustriert.

In den vorangegangenen Kapiteln ist gezeigt worden, in welcher Weise raumbezogenes Wissen vom Agenten durch das Zusammenspiel mehrerer autonomer Module erfolgt. Als Ergebnis wird die räumliche Vorstellung, die der Agent über seine Umgebung besitzt, in einer räumlichen Konfigurationsbeschreibung repräsentiert. Im nächsten Kapitel wird die Umsetzung dieses raumbezogenen Wissens in sprachliche Beschreibungen diskutiert. Eine sprachliche Beschreibung ist jedoch ebenso wie andere Verhaltensweisen im Kontext einer Situation zu betrachten.

<sup>66</sup>Dies steht im Gegensatz zur Etablierung eines sekundären Referenzsystems, bei der nur die Ausrichtung des perceptuellen Apparates gedreht wird.

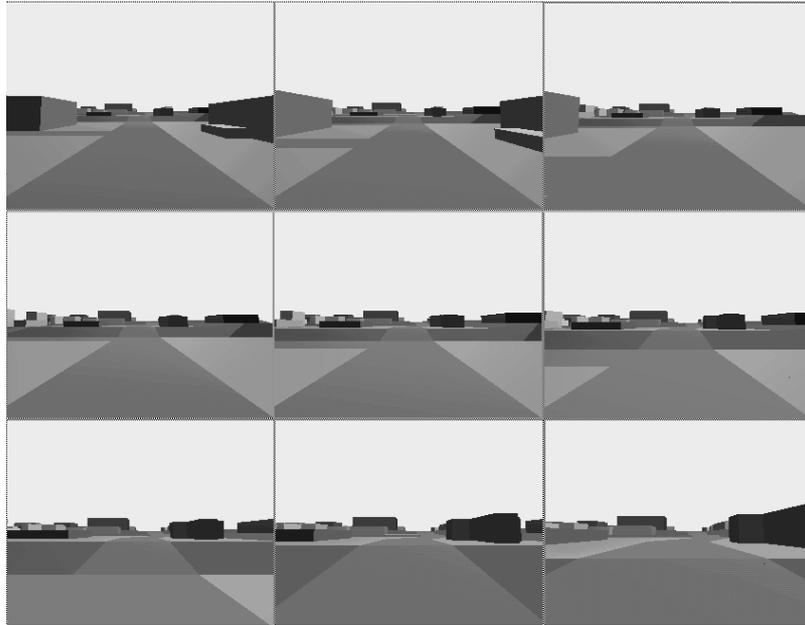


Abbildung 4.59: Schnappschüsse einer translatorischen Bewegung, die durch Evaluation von  $MOVE(v, s)$  erzeugt wurde.

#### 4.2.6 Funktionen des Sprachproduktionsmoduls

In diesem Kapitel wird die sprachbezogene generische Inhaltsstruktur vorgestellt, von der sich alle anderen Inhaltsstrukturen ableiten. Danach wird die operationale Vorgehensweise beschrieben, durch die der Inhalt einer inkrementellen Wegbeschreibung ausgewählt wird. Vor allem die Referenzierung von Landmarken wird eingehender diskutiert. Das Ergebnis dieser Funktion ist eine Oberflächenstruktur, die in sprachliche Beschreibungen (Terminalketten) transformiert wird. Abschließend wird darauf eingegangen, zu welchem Zeitpunkt eine Beschreibung gegeben wird.

Raumbezogenes Wissen ist die Grundlage für die Generierung sprachlicher Beschreibungen von Wegen. Bedingt durch die Eigenbewegung des Agenten sind neben räumlichem Wissen auch zeitliche Beschränkungen bei der Generierung zu beachten. Die Analyse inkrementeller Wegbeschreibungen (*IWB*) legt eine Fragmentierung der zugrundeliegenden Inhaltsstrukturen von fünf Inhaltskategorien (*WER-K*, *WAS-K*, *WOHIN-K*, *WO-K*, *WANN-K*) nahe (vgl. Kapitel 3). Eine aus den empirischen Untersuchungen abgeleitete *generische Inhaltsstruktur* bildet die Grundlage für die Generierung einer situationsspezifischen *IWB*. Wie der Korpus der *IWB* zeigt, lassen sich typische Kombinationen von Inhaltskategorien den Zeitrahmen zuordnen (vgl. Tabelle 3.2 in Kapitel 3.2). Eine solche Kombination von Inhaltskategorien wird im weiteren als *Inhaltsstruktur* bezeichnet. Je weniger Zeit dem Beschreiber zur Verfügung steht, um so kürzer werden die Beschreibungen und um so weniger Kategorien werden integriert.

##### 4.2.6.1 Generische Inhaltsstruktur inkrementeller Wegbeschreibungen

Aufgrund der Schematisierung der kommunikativen Handlung einer inkrementellen Wegbeschreibung werden häufig einige Kategorien, insbesondere die Kategorie *WER-K*, die das Subjekt, und *WAS-K*, die die Aktion spezifiziert, i.a. elliptifiziert und durch den Kontext als bekannt vorausgesetzt. Dies beruht darauf, daß der Rezipient sich während einer Beschreibung nicht ändert und

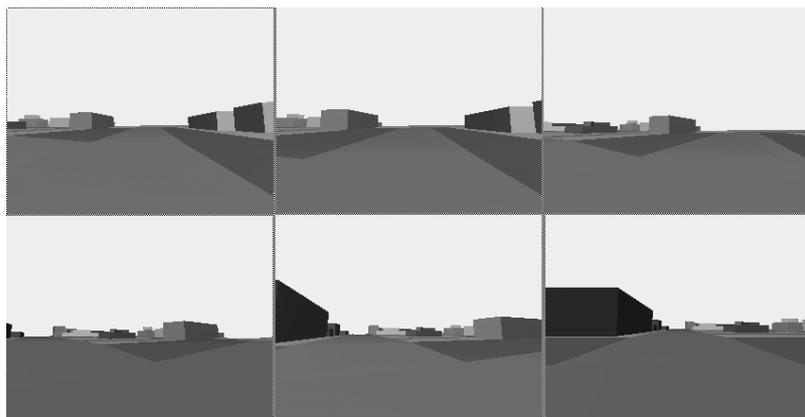


Abbildung 4.60: Schnappschüsse einer rotationalen Bewegung, die durch Evaluation von  $ROTA-TE(r, \alpha)$  erzeugt wurde.

die auszuführende Aktion durch den Bewegungsmodus und durch räumliche Angaben festgelegt ist. Kommunikative Ökonomie gebietet, daß der Rezipient nicht durch redundante Information abgelenkt wird (vgl. [Gri75]). Von besonderem Interesse sind im Kontext der *IWB* die drei anderen Kategorien *WOHIN-K*, *WO-K* und *WANN-K*. Als minimale Information benötigt der Rezipient eine Angabe der einzuschlagenden Richtung am nächsten Entscheidungspunkt (*WOHIN-K*). Die *WOHIN-K* Kategorie rekuriert auf das Wissen, welches durch *pfadbasierte Intentionen* repräsentiert wird. Weitaus komplexer ist die Instantiierung einer *WO-K* Kategorie. Sie basiert i.a. sowohl auf *Routen-* als auch auf *Landmarkenwissen* einer Situation. Im einfachsten Falle wird die Lokation einer Aktion durch einen Verzweigungspunkt indiziert. Dieser Fall wird besonders in zeitkritischen Situationen verwendet. Besitzt der Sprecher ausreichend viel Zeit, werden etwaige Ambiguitäten der Beschreibung durch Angabe von Landmarken und deren räumlichen Lagebeziehungen zu Pfadelementen vermindert. Dabei nimmt eine Landmarke die Rolle eines Referenzobjektes ein, zu der Pfadelemente in Beziehung gesetzt werden. Obschon inkrementelle Wegbeschreibung in dynamischen Umgebungen gegeben werden, sind zeitliche Referenzen rein qualitativer Natur und indizieren nur grobe zeitliche Aspekte. Zeitliche Referenzen indizieren einen deiktischen Raum und werden besonders in unmittelbarer Nähe zum Verzweigungspunkt wechselweise mit räumlichen deiktischen Referenzen verwendet.<sup>67</sup>

Wie die Analyse des Korpus ergibt (vgl. Kapitel 3), lassen sich die Beschreibungen inhaltlich durch eine *generische Inhaltsstruktur* zusammenfassen (siehe auch Abbildung 4.61), welche die Zusammenhänge zwischen Instanzen von Inhaltskategorien in einzelnen Beschreibungen repräsentiert. Von der generischen Inhaltsstruktur werden, entsprechend den zeitlichen und räumlichen Beschränkungen, *situationsbezogene Inhaltsstrukturen* abgeleitet. Eine inkrementelle Wegbeschreibung besteht im allgemeinen Falle aus einem Ereignis (EVENT), in dem der Rezipient aufgefordert wird, eine Aktion (ACTION) auf einem Pfad (PATH) an einer bestimmten Lokation ( $PLACE_{via}$ ) auszuführen.<sup>68</sup> In Fällen eines langen Pfadabschnitts wird dem Hörer eine unterstützende, *deklarative* Beschreibung gegeben, die indiziert, daß man sich immer noch auf dem richtigen Weg befindet. Die Beschreibung einer Aktion wird in einem *imparativen* Beschreibungsmodus gegeben.

Inkrementelle Wegbeschreibungen unterscheiden sich vor allem in dem Teil, welcher eine auszuführende Aktion beschreibt. In diesem Teil ist eine Instanz aus *WAS-K* ( $GO_{phys}$ ), aus *WANN-*

<sup>67</sup>Auf welchen Prinzipien der Austausch zwischen sprachlichen Referenzen wie „hier“ oder „jetzt“ erfolgt, wird nicht weiter behandelt. Erste Ansätze finden sich u.a. in [Kle79].

<sup>68</sup>Diese Notation einer Inhaltsbeschreibung entspricht der von Levelt (vgl. [Lev89]).

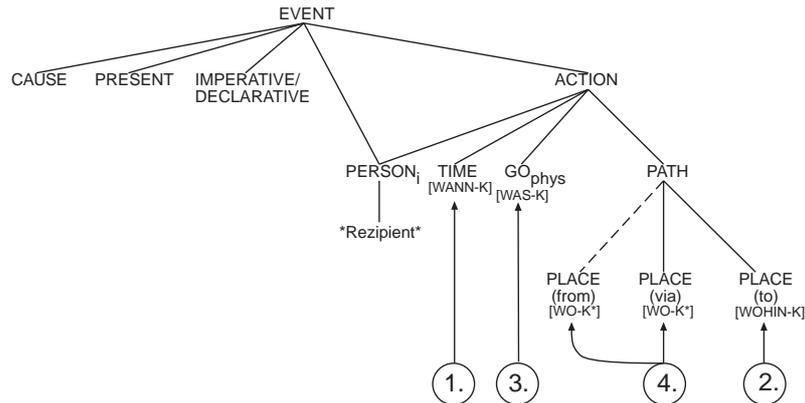


Abbildung 4.61: Generische Inhaltsstruktur für inkrementelle Wegbeschreibungen und die Abfolge der Instantiierung

K (TIME), aus WOHIN-K (PLACE<sub>to</sub>) und aus WO-K (PLACE<sub>from</sub><sup>69</sup>, PLACE<sub>via</sub><sup>70</sup>) integriert. Das „Objekt“, auf das sich diese Aktion bezieht, ist der Rezipient (\*Rezipient\*), der gleichzeitig das Subjekt der gesamten Beschreibung ist, weswegen dies eine Instanz der WER-K Kategorie ist. In den meisten inkrementellen Wegbeschreibungen fehlt eine explizite Nennung des Verbs (vgl. Satz 41).

(41) „An der Kreuzung rechts ab.“

Aber besonders im Falle eines Abbiegevorgangs werden grammatikalische Eigenschaften eines geeigneten Verbs berücksichtigt. Der vom Verb „abbiegen“ abgespaltene Partikel „ab“ indiziert alleine die gesamte Verbalphrase.<sup>71</sup> Da die Variationsmöglichkeiten einer Aktion selbst durch die Domäne eingeschränkt wird, ist die Information, die durch das Verb gegeben werden kann, redundant bzw. evident. In zeitkritischen Situationen entfällt häufig selbst der Partikel. Ein solches nicht-verbalisiertes Verb vererbt über Subkategorisierungseigenschaften Teile seiner Merkmale, wie beispielsweise den Partikel und die Möglichkeit, eine adverbiale Ergänzung aufzunehmen, an WO-K Kategorien. Die Instanziierung der generischen Inhaltsstruktur erfolgt sequentiell. Dies wird dadurch unterstützt, daß die Probanden inkrementelle Wegbeschreibungen bruchstückartig präsentierten. Dies läßt darauf schließen, daß auch die Bestimmung der inhaltlichen Struktur zumindest teilweise sequentiell erfolgt.

#### 4.2.6.2 Auswahl von Inhaltsstrukturen bezüglich einer Situation

Die Sprachgenerierung inkrementeller Wegbeschreibungen erfolgt auf der Basis einer *generischen Inhaltsstruktur* und deren Ableitungen, die durch Analyse des Korpus extrahiert worden sind (vgl. Abbildung 4.61 und Kapitel 3). Diese werden entsprechend den *zeitlichen*, *intentionalen* und *räumlichen* Gegebenheiten bezüglich einer Situation instantiiert. Die Reihenfolge, in der Kategorien instantiiert werden, strukturiert das Verhalten des Generierungsmoduls unter Zeitbeschränkung. Die funktionale Zerlegung der Generierung einer inkrementellen Wegbeschreibung wird durch die Funktion *Generate-IWB-Description* gesteuert (vgl. Funktion 11). Nachdem alle

<sup>69</sup>Referenziert die Lokation, an dem sich der Agent momentan befindet.

<sup>70</sup>Referenziert die Lokation, über die der Agent in den nächsten Pfadabschnitt gelangen möchte.

<sup>71</sup>Zusätzlich sind Verben wie „einordnen“, die ebenfalls adverbial erweitert werden können.

relevanten Kategorien instantiiert worden sind, wird die Oberflächenstruktur ausgewählt, deren Gewichtung, entsprechend den empirischen Untersuchungen (vgl. Tabelle 3.2 in Kapitel 3.2), maximal ist.

---

Funktion 11: Generate-IR-Description

```

Generate-IWB-Description( $T_{guess}$ ,  $listener_{props}$ ,  $T_{old}$ )
  content-structure-set := Determine-Content-Structure-Set( $T_{guess}$ ,  $listener_{props}$ ,  $T_{old}$ );
  transformation-set := Generate-WHEN-K(transformation-set);
  transformation-set := Generate-WHERE-TO-K(transformation-set);
  transformation-set := Generate-WHAT-K(transformation-set);
  transformation-set := Generate-WHERE-VIA-K(transformation-set);
  surface-structure := Select-Most-Specified(transformation-set);
  return(surface-structure);

```

---

Beruhet eine Beschreibung auf der gesamten generischen Inhaltsstruktur, erfolgt die Transformation der Inhaltsstruktur des Aktionsteils in eine Oberflächenstruktur in vier Schritten. Vom Zeitrahmen  $T_{guess}$  ausgehend, wird in einem ersten Schritt eine natürlichsprachliche Beschreibung einer zeitlichen Referenz abgeleitet (WANN-K, 1. Schritt in Abbildung 4.61). Dieser Evaluation nachgeordnet ist die zur Instantiierung von *WOHIN-K*, welches die Orientierung einer Aktion indiziert (2. Schritt). Die zugrundeliegende Richtungsinformation wird von den pfadbasierten Intentionen abgeleitet. Die Etablierung einer geeigneten Instanz der *WOHIN-K* Kategorie initiiert die Auswahl einer *WAS-K* Instanz (3. Schritt). Entweder handelt es sich um einen Abbiegevorgang oder um eine Fortsetzung der Bewegung in Geradeausrichtung. Falls zeitliche Beschränkungen es zulassen, der Rezipient es erwünscht und geeignetes Landmarkenwissen vorhanden ist, wird in einem vierten Schritt die Lokation, an der die Aktion stattfinden soll, genauer spezifiziert. Eine Lokation kann durch unterschiedlich viele Landmarken und deren Eigenschaften beschrieben werden, was durch eine Sequentialisierung einzelner *WO-K* Instanzen erreicht wird.<sup>72</sup> In seltenen Fällen wird auch die aktuelle Lokation als Referenzpunkt eingeführt ( $PLACE_{from}$ ). Die Transformation einer Instanz einer solchen Kategorie erfolgt analog zur *WO-K* Instanz, die den Verzweigungspunkt indiziert ( $PLACE_{via}$ ).

**4.2.6.2.1 Auswahl von situationsbezogenen Inhaltsstrukturen** Zeitliche Beschränkungen beziehen sich auf das Verhältnis zwischen dem Agenten, dem Rezipienten und der Umgebung. Alle Beschreibungen, die der Hörer aufgrund ihrer Komplexität nicht bis zum Erreichen des Verzweigungspunktes verstehen und umsetzen kann, gelten als nicht adäquat.<sup>73</sup> Wie aus dem Korpus hervorgeht (siehe Analyse der zeitlichen Kategorien in Kapitel 3), lassen sich fünf Zeitrahmen unterscheiden, in denen die Probanden unterschiedlich strukturierte Beschreibungen gegeben haben (vgl. Kapitel 3, 4.2.1). Die Annahme bei der Bestimmung einer Beschreibung lautet, daß ein Sprecher immer versucht, in einer gegebenen Situation eine möglichst gute Beschreibung zu generieren. Hierzu dient die Heuristik, daß je genauer die Information, um so besser die jeweilige Beschreibung ist. Über das Zeitintervall  $T_{guess}$  werden a priori situationsabhängig Inhaltsstrukturen von der generischen Inhaltsstruktur abgeleitet.

Hierzu berechnet die Funktion *Determine-Content-Structure-Set* (vgl. Funktion 12), mit welchem Zeitrahmen der aktuelle Zeitrahmen  $T_{guess}$  übereinstimmt (*Match-Current-Time-Frame*). Nicht in jedem Fall wird dem Hörer eine maximal detaillierte Beschreibung gegeben. Ist der

---

<sup>72</sup>Diese Iterationsmöglichkeit ist in der generischen Inhaltsstruktur durch einen Stern markiert: WO\*.

<sup>73</sup>Ein solches Konzept der *Adäquatheit* gilt nicht für eine dialogische Situation. In einem Dialog kann der Rezipient im Zweifelsfalle nachfragen, bzw. sich rückversichern.

## Funktion 12: Determine-Content-Structure-Set

```

Determine-Content-Structure-Set( $T_{guess}$ ,  $listener_{props}$ ,  $T_{old}$ )
   $T_{current}$  := Match-Current-Time-Frame( $T_{guess}$ );
   $T_{listener}$  := Match-Listener-Time-Frame( $listener_{props}$ )
   $T_{select}$  := min( $T_{current}$ ,  $T_{listener}$ ,  $T_{old}$ );
  content-structure-set := Select-Structure-Schema( $T_{select}$ );

```

Hörer an einer möglichst detaillierten Beschreibung interessiert, so wird durch die Funktion *Match-Listener-Time-Frame* der Wert von  $T_{listener}$  auf  $\infty$  gesetzt, so daß der maximal mögliche Zeitrahmen durch  $T_{current}$  vorgegeben wird. Ist zu der aktuellen Situation bereits eine Beschreibung im Zeitrahmen  $T_{old}$  gegeben worden, so ist dies für die nachfolgende Beschreibung zu beachten. Eine nachfolgende Beschreibung kann dann erst im übernächsten Zeitrahmen ( $T_{old} + 2$ ) gegeben werden. Ist  $T_{listener}$  ungleich  $\infty$ , so ist der maximale Zeitrahmen  $T_{select}$ , der für die Beschreibung zur Verfügung steht, durch das Minimum über die Zeitrahmen  $T_{guess}$ ,  $T_{listener}$  und  $T_{old} + 2$  bestimmt. Dieses spiegelt wider, daß der Sprecher den Rezipienten nicht unstrukturiert mit Information überfordert, sondern eine Grobstruktur verwendet, die nur zu bestimmten Zeitrahmen Beschreibungen ermöglicht. Das Zeitintervall  $T_{guess}$  gibt vor, zu welchem Zeitrahmen  $T_i$  die nächste Beschreibung potentiell generiert werden kann (vgl. Kapitel 4.2.1). Jedem der fünf Zeitrahmen lassen sich Ableitungen der generischen Inhaltsstruktur zuordnen, die unterschiedliche Kombinationen von *W\*-Kategorien* enthalten (vgl. Abbildung 4.62). Je näher dieser Zeitrahmen vor dem Erreichen des Verzweigungspunktes ist, um so kürzer ist die resultierende Beschreibung. Von den beiden Zeitrahmen  $T_{current}$  und  $T_{listener}$  wird derjenige, der am nächsten zum Verzweigungspunkt ist, ausgewählt. Bezüglich dieses Zeitrahmens bestimmt sich die Menge der möglichen Inhaltsstrukturen (siehe Tabelle 3.2 in Kapitel 3.2 und Abbildung 4.62). Tabelle 3.2 zeigt eine Zuordnung der einzelnen Kategorien zu Zeitrahmen, wobei jeder einzelnen Zeile eine Konjunktion von Inhaltskategorien zugeordnet ist, die in Abbildung 4.62 graphisch illustriert ist. Die Gewichtung gibt an, mit welcher Häufigkeit eine situationsabhängige Inhaltsbeschreibung bezüglich eines Zeitrahmens verwendet wird (vgl. Kapitel 3).

Wenn die Menge der Inhaltsstrukturen für die aktuelle Situation ausgewählt worden ist, werden die Kategorien der einzelnen Inhaltsstrukturen in Oberflächenstrukturen transformiert. Der Rezipient ist in allen Fällen konstant und wird optional durch eine geeignete Nominalphrase gefüllt. Danach werden zeitbezogene, gefolgt von intentionsbezogenen und abschließend lokationsbezogenen Kategorien im Wissen des Agenten über die Situation verankert. Durch diese Schritte erfolgt die Transformation von Inhaltsstrukturen in Oberflächenstrukturen. Bei der Instantiierung einer Inhaltskategorie wird die Wortwahl getroffen und dadurch, entsprechend Levelts Hypothese ([Lev89]), die Phrasenstruktur der Instanz über grammatikalische Regeln an die gesamte Struktur vererbt.

**4.2.6.2.2 Zeitliche Referenzierung** Zeitliche Referenzen auf sprachlicher Ebene korrespondieren eng mit solchen auf der Steuerungsebene. Über die Größe von  $T_{guess}$  entscheidet die Funktion *Generate-WHEN-Category* (vgl. Funktion 13), ob und welches Lexem eine geeignete Beschreibung der zeitlichen Zusammenhänge zwischen der Bewegung des Agenten und der Situation repräsentiert.

Wie der Tabelle 3.2 in Kapitel 3.2 entnommen werden kann, enthalten verschiedene Inhaltsstrukturen Kategorien des Typ *WANN-K*. Geeignete Lexeme werden durch die Aktivatoren @now, @soon, @later ausgewählt, die zu einer Adverbialphrase transformiert werden. Dabei wird die

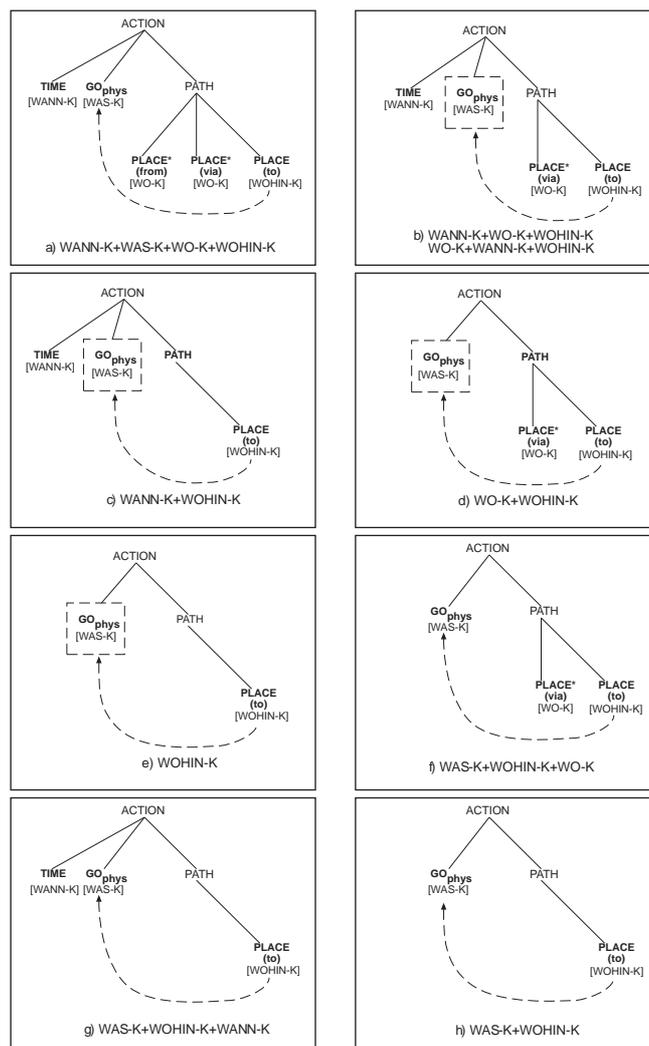


Abbildung 4.62: Von der generischen Inhaltsstruktur abgeleitete Strukturen des ACTION-Teils

adverbiale Ergänzung des Verbs um die der zeitlichen Referenz erweitert. Die Funktion *Lexicalize* ersetzt in der Inhaltsstruktur einen TIME-Marker, der durch die ersten Parameter gegeben ist, durch eine Phrasenstruktur, die dem Lexem in der zweiten Position zugeordnet ist.

**4.2.6.2.3 Instantiierung einer Richtungs- und Aktionsreferenz** Die Bedeutung einer WOHIN-K Kategorie ist es, die Orientierung der am nächsten Verzweigungspunkt auszuführenden Aktion zu indizieren (vgl. Funktion 14).

Entsprechend der pfadbasierten Intention mit dem höchsten Anwendbarkeitsgrad  $AG_{int}$  bzw.  $AG_{kpi}$  wird eine geeignete Adverbialphrase ausgewählt, die die Instanz der WOHIN-K Kategorie ersetzt. Die pfadbasierte Intention hat Auswirkungen auf die Transformation der WAS-K Instanz. Handelt es sich um einen Abbiegevorgang, d.h., daß die Intention  $\$turn-left\$$  oder  $\$turn-right\$$  enthält, wird durchweg eine Verbphrase etabliert, welche eine solche Aktion beschreibt (vgl. Funktion 15).

**4.2.6.2.4 Integration von räumlichem Landmarken- und Routenwissen** Der Hauptbestandteil einer inkrementellen Wegbeschreibung bezieht sich auf eine elaborierte Verwendung

---

 Funktion 13: Generate-WHEN-Category
 

---

*Generate-WHEN-Category(transformation-set)*  
 Case  $T_{\text{select}}$  is equal  
 $T_5$ : Lexicalize(TIME, @now)  
 $T_4$ : Lexicalize(TIME, @soon)  
 $T_3$ : Lexicalize(TIME, @later)

---



---

 Funktion 14: Generate-WHERE-TO-Category
 

---

*Generate-WHERE-TO-Category(transformation-set)*  
 If #right-of#(ep, p<sub>2</sub>) then  
     Lexicalize(PLACE-TO, #right#, nil)  
 If #left-of#(ep, p<sub>2</sub>) then  
     Lexicalize(PLACE-TO, #left#, nil)  
 If #behind#(ep, p<sub>2</sub>) then  
     Lexicalize(PLACE-TO, #straight#, nil)

---

der Kategorie *WO-K*. Die Beschreibung eines Gebietes, an dem eine Aktion auszuführen ist, erfolgt entweder durch Indikation einer oder mehrerer Landmarken oder durch Referenzierung des Gebiets, welches durch den Verzweigungspunkt und den nachfolgenden Pfadabschnitt konstituiert wird (vgl. Funktion 16). Dazu wird zuerst durch die Funktion *Select-AppL-LM-Set* die Menge aller Landmarken ausgewählt, deren Anwendbarkeitsgrad maximal ist (vgl. Funktion 17).

**4.2.6.2.4.1 Referenzierung von Landmarken** Eine Landmarke wird dann in die Beschreibung integriert, wenn sie sich in einer geeigneten räumlichen Lokation relativ zum aktuellen Pfadabschnitt  $p_{\text{current}}$  zum nächsten Pfadabschnitt  $p_{\text{next}}$  und zum Verzweigungspunkt  $ep$  befindet (vgl. Funktion 18). Ob eine Landmarke als Referenz für die Beschreibung einer Aktion verwendbar ist, hängt vom *räumlichen Anwendbarkeitsgrad* einer Landmarke (RAG) relativ zu den Pfadelementen als auch von seiner *visuellen Saliens* (VS) ab. Allgemein läßt sich der Anwendbarkeitsgrad (AG) durch eine Verknüpfung dieser beiden Eigenschaften einer Landmarke durch zwei Funktionen  $F$  und  $G$  angeben (vgl. Gleichung 42).

$$(42) \quad \text{AG}(lm) = F(\text{RAG}(lm)) \otimes G(\text{VS}(lm))$$

Aufgrund der Einfachheit ist eine Linearkombination eine erste Approximation des Operators  $\otimes$ . Bezogen auf das Experiment zur Untersuchung der visuellen Saliens hat sich die Identitätsfunktion als geeignete Modellierung für  $F$  herausgestellt (vgl. Kapitel 3.3). Die Funktion  $G$  bezieht sich auf die visuelle Saliens (VS( $lm$ )) einer Landmarke (vgl. Kapitel 4.2.3). Hierüber ist repräsentiert, ob die Landmarke visuell von anderen Landmarken unterscheidbar ist und somit maximal disambiguierend wirkt (vgl. Kapitel 4.2.3). Visuelle Saliens haben einen stärkeren Einfluß auf die Anwendbarkeit einer Landmarke, so daß die Euler-Funktion eine geeignete Charakteristik bietet. Daraus ergibt sich für den Anwendbarkeitsgrad einer Landmarke eine Linearkombination aus räumlicher Anwendbarkeit und visueller Saliens (vgl. Gleichung 43).

$$(43) \quad \text{AG}(lm) = \text{RAG}(lm) + e^{\text{VS}(lm)},$$

---

 Funktion 15: Generate-WHAT-Category
 

---

```

Generate-WHAT-Category(transformation-set)
  If #right-of#(ep, p2) ∨ #left-of#(ep, p2) then
    Lexicalize(GO, turn, nil)
  If #behind#(ep, p2) then
    Lexicalize(GO, go, nil)
  
```

---



---

 Funktion 16: Generate-WHERE-VIA-Category
 

---

```

Generate-WHERE-VIA-Category(transformation-set)
  lm-set := Select-Appl-LM-Set;
  if lm-set ≠ nil then
    Descr-Location-By-LM(lm-set)
  else
    Descr-Location-By-Path-Elements
  
```

---

Nicht in jedem Fall reicht der maximale Anwendbarkeitsgrad einer Landmarke aus, um sie als Referenzobjekt verwenden zu können. Besitzt eine Landmarke in einer Situation zwar den höchsten Anwendbarkeitsgrad AG, aber ist dieser zu gering, so wird die Landmarke nicht als Referenzobjekt verwendet.<sup>74</sup> Als Rückgabewert liefert in diesem Fall die Funktion *Appl-Degree* den Wert  $-\infty$ . Im anderen Fall berechnet sie den Anwendbarkeitsgrad der Landmarke bezüglich der aktuellen Situation (vgl. Funktion 18).

In Abbildung 4.63 ist ein Beispiel gezeigt, in der die Landmarke *A* einen maximalen Anwendbarkeitsgrad besitzt, der aber nicht ausreicht, um *A* aufgrund visueller Merkmale als Referenzobjekt verwenden zu können.

Die in einer Situation als Referenzobjekte anwendbar klassifizierten Landmarken werden in der Menge *lm-set* gesammelt (*Generate-WHERE-VIA-Category*). Ist *lm-set* nicht leer, so wird die intendierte Aktion am nächsten Verzweigungspunkt durch diese Landmarken referenziert (vgl. Funktion 19). Ansonsten wird die Lokation durch Pfadelemente, und hierbei besonders durch die Lokation des Verzweigungspunktes indiziert (vgl. Funktion 23).

In den meisten Fällen bezogen sich die Probanden auf genau eine Landmarke (vgl. Kapitel 3.2). Bevor eine Landmarke beschrieben wird, muß überprüft werden, ob ihre Lokation mit denen der Pfadelemente konsistent ist (vgl. Funktion 20). Wie aus dem Korpus hervorgeht, werden keine Landmarken beschrieben, die bei einem Abbiegevorgang auf der gegenüberliegenden Halbebene liegen (vgl. Kapitel 3.2). Wie die Untersuchungen ferner zeigen, wird dies zusätzlich dadurch gestützt, daß es den Probanden erhebliche Mühe bereitete eine Landmarke auf der gegenüberliegenden Halbebene in ihrer Beschreibung zu integrieren.

Bei der Instantiierung der *WO-K* Kategorie lassen sich drei Fälle unterscheiden. Liegt eine Abbiegeintention vor, so kann der Verzweigungspunkt und der nachfolgende Pfadabschnitt entweder „hinter“ oder „vor“ der zu referenzierenden Landmarke liegen. Der dritte Fall betrifft

---

<sup>74</sup>Daß nur solche Landmarken als Referenzobjekte verwendet werden können, die einen maximalen Anwendbarkeitsgrad besitzen (vgl. Funktion *Select-Appl-LM-Set* in Funktion 17), kann dahingehend abgeschwächt werden, daß auch Landmarken berücksichtigt werden, deren Anwendbarkeitsgrad in einem Bereich unterhalb des Maximums der Situation liegen. Besonders im Hinblick auf eine Verwendung der Präposition „zwischen“ ist eine solche Vorgehensweise effektiv.

## Funktion 17: Select-Appl-LM-Set

```

Select-Appl-LM-Set
max-appl-degree := 0;
for all lm ∈ Salient-Landmarks(rkb) do
  ad := Appl-Degree(lm);
  if ag == max-appl-degree then
    integrate-lm-set := union(lm, integrate-lm-set)
  if ag > max-appl-degree then
    max-appl-degree := ad;
    integrate-lm-set := {lm}
return(integrate-lm-set);

```

## Funktion 18: Appl-Degree(lm)

```

Appl-Degree(lm)
if AG(lm) ≥ appl-degree-threshold then
  return(appl-degree)
else
  return(-∞)

```

eine Intention der Fortbewegung *geradeaus* in Frontalrichtung, wobei keine Landmarken als Referenzobjekte verwendet werden (vgl. Abbildung 4.64b).

Die ersten beiden Fälle der Referenzierung durch eine Landmarke verlaufen für Intentionen des Linksabbiegens symmetrisch zu solchen des Rechtsabbiegens, weswegen nur der erstere weiter betrachtet werden soll (vgl. Abbildung 4.64a). Bei der Auswahl einer geeigneten räumlichen Beschreibung der Lokation einer zu referenzierenden Landmarke werden ausschließlich intrinsische Relationen zwischen der Landmarke, dem Verzweigungspunkt (C) und dem nachfolgenden Pfadabschnitt (S2) berücksichtigt. Die Auswahl einer räumlichen Präposition erfolgt über eine Evaluation solcher intrinsischer Relationen und C bzw. S2, die im Anwendbarkeitsmuster einer räumlichen Präposition (*PAGM*) angegeben sind. Zur Illustration der verwendeten Methode werden im weiteren Verlauf o.B.d.A. die räumlichen Präpositionen „hinter“, „vor“ und „zwischen“ diskutiert (vgl. Abbildung 4.65).

Die Anwendbarkeit (*AGP*) einer räumlichen Präposition (*sp-prep*) wird durch das arithmetische Mittel der quadrierten Anwendbarkeitsgrade der räumlichen Relationen der assoziierten Landmarke bestimmt (vgl. Definition 44).

$$(44) \quad AGP(sp-prep) = 1/n \sum_{i=1}^n rrel_i(PAGM(prepare, lm))$$

Der Anwendbarkeitsgrad einer räumlichen Präposition mit dem höchsten Wert wird zur Beschreibung der räumlichen Lage des Gebietes, in dem die Aktion auszuführen ist, und der referenzierenden Landmarke verwendet. Die Bestimmung einer AGP erfolgt durch Vergleich der intrinsischen Relationen der räumlichen Konfigurationsbeschreibung bezüglich der Landmarke mit den geforderten Relationen der assoziierten PAGM. Die Auswertung erfolgt in zwei Schritten. Zuerst wird überprüft, ob alle vom PAGM geforderten räumlichen Relationen zwischen der Landmarke und den Pfadelementen bestehen. Ist dies nicht der Fall, scheidet die Auswertung. Bestehen alle Relationen, so werden sie entsprechend der Funktion  $AGP(sp-prep)$  ausgewertet. Die räumliche

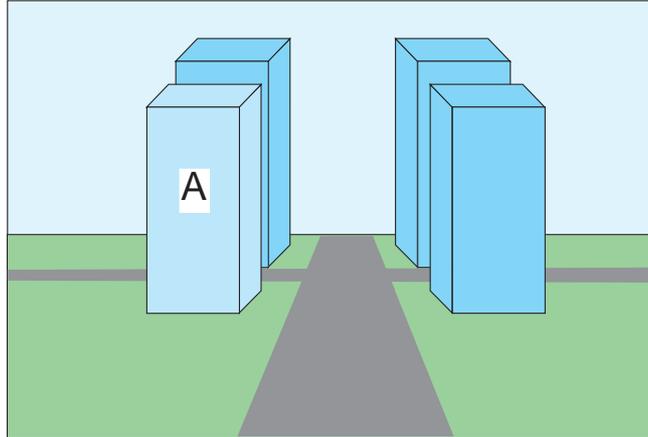


Abbildung 4.63: Beispiel für eine Situation, in der eine Landmarke A eine maximale visuelle Salienz besitzt, die aber relativ zu den anderen Objekten für eine Referenzierung zu gering ist

	S1	C	S2
LM <sub>A</sub>	#right-of# = 0,8 #in-front-of# = 0,6	#right-of# = 0,7 #behind# = 0,7	#right-of# = 0,9

Tabelle 4.4: Intrinsische Relationen und Anwendbarkeitsgrade zwischen der Landmarke A und relevanten Pfadelementen

Präposition mit dem maximalen Anwendbarkeitsgrad wird letztendlich ausgewählt.

In Abbildung 4.64a ist eine Beispielsituation, in der A die zu referenzierende Landmarke ist. Geht man davon aus, daß zwischen A und den beiden Pfadelementen C und S2 alle notwendigen räumlichen Relationen wie in Tabelle 4.4 bestehen, so ergibt sich ein maximaler PAGM bezüglich einer pfadbasierten Intention  $\$turn-left\$$  für die räumliche Präposition „hinter“ (vgl. Berechnung 45). Die Auswertung der PAGM(vor) scheitert aufgrund der fehlenden #in-front-of# Relation zur S2 und C.

$$(45) \quad \text{PAGM}(\text{hinter}, A) = (0,9^2 + 0,7^2 + 0,7^2)/3 = 0,45$$

Über eine Funktion *Det-Max-PAGM* wird die räumliche Präposition *spatial-prep* mit dem maximalen Anwendbarkeitsgrad *AGP* bestimmt. Liegt der AGP dieser räumlichen Präposition über einem zugeordneten Schwellwert *Threshold(spatial-prep)*, so wird diese Präposition dazu verwendet, die Lokation der Aktion bezüglich der referenzierenden Landmarke zu beschreiben (vgl. Funktion 19)

Der dritte Parameter der Funktion *Lexicalize* enthält den rekursiven Anteil, der durch die angegebene Funktion weiter berechnet wird. In Abbildung 19 ist dies Descr-LM(lm). Das Ergebnis der Auswertung wird von *Lexicalize* mit dem Wert des zweiten Parameters, in diesem Falle *spatial-prep*, unifiziert.

Die Beschreibung einer Landmarke selbst wird durch die Funktion *Descr-LM* bestimmt (vgl. Funktion 21). Sie bestimmt zu einer gegebenen Landmarke geeignete Merkmale, die es dem Rezipienten erlauben, diese Landmarke ohne Ambiguitäten zu identifizieren.<sup>75</sup> Über die Men-

<sup>75</sup>Im Rahmen dieser Arbeit handelt es sich dabei ausschließlich um physikalisch-visuelle Merkmale.



lo refo	S2	C
LM	#behind#	#behind# #right-of#

PAGM für „hinter“  
bei \$turn-left\$  
a)

lo refo	S2	C
LM	#in-front-of#	#in-front-of# #right-of#

PAGM für „vor“  
bei \$turn-left\$  
b)

lo refo	S2	C
LM	#behind#	#behind# #right-of#

PAGM für „hinter“  
bei \$turn-right\$  
c)

lo refo	S2	C
LM	#in-front-of#	#in-front-of# #right-of#

PAGM für „vor“  
bei \$turn-right\$  
d)

lo refo	S2	C
LM <sub>1</sub>	#right-of#	#right-of# #behind#
LM <sub>2</sub>	#left-of#	#left-of# #behind#

PAGM für „zwischen“  
bei \$go-straight\$  
e)

lo refo	S2	C
LM <sub>1</sub>	#right-of#	#right-of# #in-front-of#
LM <sub>2</sub>	#left-of#	#left-of# #in-front-of#

PAGM für „zwischen“  
bei \$go-straight\$  
f)

Abbildung 4.65: Anwendbarkeitsmuster PAGM der räumlichen Präpositionen „vor“, „hinter“ und „zwischen“ bei unterschiedlichen pfadbezogenen Intentionen

Die Landmarke und die Merkmale werden durch eine Adjektivphrase lexikalisiert.

**4.2.6.2.4.2 Referenzierung von Pfadelementen** Existieren keine Landmarken mit einem ausreichenden hohen Anwendbarkeitsgrad, werden Pfadelemente dazu verwendet, das Gebiet der auszuführenden Aktion genauer zu spezifizieren. Eine Analyse des Korpus ergibt, daß die Beschreibung eines Gebietes der auszuführenden Aktion, basierend auf Pfadelementen, grundsätzlich durch eine Referenzierung des Verzweigungspunktes unter Verwendung der räumlichen Präposition „an“ erfolgt.

Besitzt der Verzweigungspunkt visuell-saliente Merkmale, werden diese zur genaueren Beschreibung verwendet (*Descr-Path-Element*, analog zu *Descr-LM*). In dieser werden auffällige Eigenschaften des Verzweigungspunktes extrahiert und durch die Funktion *Lexicalize* in Oberflächenstrukturen transformiert.<sup>77</sup>

**4.2.6.2.4.3 Lexikalisierung und grammatikalische Regeln inkrementeller Wegbeschreibungen** Der Zugriff auf das Lexikon, die Wortwahl und der Aufbau einer syntaktischen

tionen mit der steigenden Anzahl von Merkmalstypen.

<sup>77</sup>Eigenschaften von Verzweigungspunkten sind im Rahmen dieser Arbeit nicht näher untersucht worden. Gerade für Verzweigungspunkte ist zu vermuten, daß semantische Faktoren eine entscheidende Rolle spielen.

## Funktion 20: Spatially-Consistent

```

Spatially-Consistent(lm, p1, p2, ep)
  if spatially-near(lm, p2) then
    if Type(SpRel(lm, p2)) == Type(SpRel(p2, ep)) then
      return(t);
    if #behind#(ep, p2) ≥ Threshold(#behind#) then
      return(t);
  else;
  return(nil);

```

## Funktion 21: Descr-LM

```

Descr-LM(lm-set)
  For all l ∈ All-Selected-LMs ∧ l not ∈ lm-set do
    f-setl := Saliient-Features(l);
    f-set := union(f-set, f-setl)
  f-setlm := Saliient-Features(lm);
  distinct-f-set := Select-Distinctive-Features(f-setlm, f-set);
  Lexicalize(PLACEvia, (first distinct-f-set), ((rest distinct-f-set), lm))

```

Teilphrase erfolgt durch die Funktion *Lexicalize* (vgl. Funktion 24). In *Lexicalize* sind die grammatikalischen Regeln kodiert, die durch Analyse des Korpus abgeleitet worden sind. Tabelle 4.5 enthält die aus dem Korpus abgeleiteten grammatikalischen Regeln inkrementeller Wegbeschreibungen. Als ersten Parameter übernimmt *Lexicalize* einen Terminalknoten einer Inhaltsstruktur, welcher entsprechend dem zweiten und dritten Parameter transformiert werden soll. Der zweite Parameter trägt den Wert des Phrasenkopfes, entsprechend dessen Typs die geeignete Phrasenstruktur ausgewählt wird. Der dritte Parameter enthält den Teil, der dieser Phrase untergeordnet ist. Durch rekursiven Aufruf wird eine Ersetzung dieses Teils erreicht. Der durch den dritten Parameter entstandene Teilbaum der Oberflächenstruktur wird an die geeignete Stelle der Phrase eingehängt. Strukturelle Merkmale werden entsprechend unifikationsgrammatikalische Techniken vererbt (*Generate-Surf-Struct*). Eine genauere Analyse dieser Vererbungsmechanismen liegt außerhalb des Fokus dieser Arbeit (für verschiedene Ansätze siehe u.a. [Kay79], [Usz86]).<sup>78</sup> Der Rückgabewert der Funktion ist eine instantiierte Phrasenstruktur *surfstruct*.

Die aus dem Korpus abgeleiteten Regeln der zugrundeliegenden Grammatik enthalten als Terminale Verben (V), Partikelabspaltungen vom Verb (Prt), Artikel (Det) und Adverbien (Adv). Diese Terminale können zu Verbphrasen (VP, VP')<sup>79</sup>, Nominalphrasen (NP, NP'), Adjektivphrasen (AdjP, AdjP'), Präpositionalphrasen (PP, PP') und Adverbialphrasen (AdvP, AdvP') zusammengesetzt werden. Ein Satz S leitet sich im Falle inkrementeller Wegbeschreibungen direkt von einer Verbphrase (VP') ab.

Gibt es nach Evaluation dieser Funktion alternative Oberflächenstrukturen zur Beschreibung der Situation, wird unter diesen eine solche ausgewählt, die einerseits vollständig instantiiert ist und die andererseits eine maximale Gewichtung entsprechend Tabelle 3.2 in Kapitel 3.2 besitzt (Funktion *Select-Most-Specified*).

<sup>78</sup>In der Implementation wird u.a. ein TAG-Generator (vgl. [Jos85], [HFS91], [KF95]) verwendet (vgl. [Pau95]).

<sup>79</sup>Nach der  $\bar{X}$ -Grammatik von Jackendoff (vgl. [Jac72]) werden Regeln nicht beliebig rekursiv, sondern nur entsprechend einer fest vorgegebenen Tiefe evaluiert. Eine solche Tiefe ist für jede Phrase durch die Anzahl der Hochkommata angegeben.

## Funktion 22: Select-Distinctive-Features

```

Select-Distinctive-Features(f-setlm, f-set)
  dist-f-setlm := nil
  For all feature ∈ f-setlm do
    If Feature-Distance(feature, f) ≥ Threshold(F-Type(feature)) then
      dist-f-setlm := union(dist-f-setlm, feature)
      f-counter++;
    If f-counter ≥ max-distinct-features then
      return(dist-f-setlm)
  done
  return(dist-f-setlm)

```

## Funktion 23: Descr-Location-By-Path-Elements

```

Descr-Location-By-Path-Elements
  Lexicalize(PLACEvia, 'an', (Descr-Path-Element(ep)))

```

#### 4.2.6.3 Beispiele der Verwendung räumlicher Konfigurationsbeschreibungen zur Generierung inkrementeller Wegbeschreibungen

In Abbildung 4.66 sind drei Beispiele gezeigt, die die Intention, in einer identischen Situation rechts abbiegen zu wollen, durch unterschiedliche inkrementelle Wegbeschreibungen ausdrücken. Das erste Beispiel verwendet eine *WO-K+WANN-K+WOHIN-K* Inhaltsstruktur. Da die assoziierte *minimale räumliche Konfigurationsbeschreibung* keine geeignete Landmarke enthält, wird die Lokation der Aktion alleinig durch Nennung der Kreuzung indiziert. Im zweiten Beispiel wird eine *WANN-K+WAS-K+WO-K+WOHIN-K* Inhaltsstruktur verwendet. Die assoziierte *erweiterte räumliche Konfigurationsbeschreibung* enthält zwei Landmarken, von denen *L4* den höchsten räumlichen Anwendbarkeitsgrad besitzt. Aus diesem Grund wird sie in der Beschreibung durch Referenzierung des „grauen Hauses“ integriert. Eine der kürzesten Beschreibungen ist durch die häufig verwendete *WO-K+WOHIN-K* Inhaltsstruktur möglich. Auch in diesem Falle enthält die *minimale räumliche Konfigurationsbeschreibung* keine Landmarke, so daß wiederum auf die Lokation der Kreuzung referenziert wird.

#### 4.2.6.4 Der Präsentationsspeicher

Inkrementelle Wegbeschreibungen sind nicht nur inhaltlich schematisiert, sondern unterliegen auch starken zeitlichen Vorgaben. Ein Proband beschrieb diesen Umstand wie folgt: „Ich könnte jetzt schon sagen, wohin wir gehen müssen, aber es ist noch zu früh.“ Modelliert wird dieses Verhalten dadurch, daß Beschreibungen bezüglich eines bestimmten Zeitrahmens generiert werden, über den bestimmt wird, *wann* die Präsentation erfolgt.

Bevor eine neue Beschreibung im Präsentationsspeicher übernommen wird, wird der Speicher initialisiert um die bisherige Beschreibung zu deaktivieren (Initialize-Presentation-Buffer). Die Repräsentation der Beschreibung enthält die Information über den Zeitrahmen *ts*. Diese temporale Information dient dem Präsentationsspeicher dazu, den Zeitpunkt zu bestimmen, an dem die Beschreibung in der Situation gegeben werden soll.

Die Notwendigkeit für die Autonomie des mit der Funktion *Activate-Presentation-Buffer* assoziierten Prozesses ist stark durch die Ressourcenadaptivität des Gesamtmodells vorgegeben (vgl.

$S$	$\rightarrow VP'$
$VP'$	$\rightarrow (\mathbf{V}) - (NP) - (\mathbf{Prt}) - \left( \begin{array}{c} PP \\ AdvP \\ NP \end{array} \right) - (VP)$
$VP$	$\rightarrow (\mathbf{Prt}) - \mathbf{V}$
$NP'$	$\rightarrow (\mathbf{Det}) - (AdjP') - \mathbf{N} - (NP) - \left( \begin{array}{c} PP \\ VP \end{array} \right)$
$NP'$	$\rightarrow (\mathbf{Det}) - (AdjP') - \mathbf{N}$
$AdjP'$	$\rightarrow \mathbf{Adj} - (AdjP)$
$AdjP$	$\rightarrow (\mathbf{Prt}) - \mathbf{Adj} - (AdjP)$
$PP'$	$\rightarrow \mathbf{Prp} - NP' - \left( \begin{array}{c} NP \\ PP \\ VP \end{array} \right)$
$PP$	$\rightarrow \mathbf{Prt} - NP'$
$AdvP'$	$\rightarrow (AdvP) - \mathbf{Adv}$
$AdvP$	$\rightarrow \mathbf{Adv}$

Tabelle 4.5: Grammatikalische Regeln inkrementeller Wegbeschreibungen

Funktion 25). In jeder Situation bestimmt der Agent zuerst eine *minimale Beschreibung*, die er anschließend schrittweise zu *erweiterten Beschreibungen* ausbaut, falls der zeitliche Rahmen ausreichend ist. Diese Erweiterung von Beschreibungen wird mit der initialen Generierung verzahnt aktiviert (siehe Kapitel 4.2.1). Solange eine Erweiterung der aktuellen Beschreibung noch nicht vollständig generiert worden ist, gibt der Präsentationsprozeß im aktuellen Zeitrahmen, die im Präsentationsspeicher abgelegte Beschreibung aus (*Present-Description*). Nachdem eine Beschreibung gegeben worden ist, wird der Präsentationsspeicher initialisiert.

## Funktion 24: Lexicalize

```

Lexicalize(c-leaf, s-node, rec-node)
lex := Select-Lexem(s-node);
Case c-leaf is equal to
  TIME:
    Case Type(s-node) is equal
      temporal-ref: p-type := AdvP;
  GO:
    Case Type(s-node) is equal
      go: p-type := VP;
  PLACE-FROM:
    Case Type(s-node) is equal
      deictic-ref: p-type := AdvP;
  PLACE-VIA:
    Case Type(s-node) is equal
      sp-rel: p-type := PP;
      iota: p-type := NP;
      feature: p-type := AP;
      object: p-type := NP;
  PLACE-TO:
    Case Type(s-node) is equal
      sp-rel: p-type := AdvP;
      feature: p-type := AP;
      object: p-type := NP;
  nil:
    return(nil)
surfstruct := Generate-Surf-Struct(p-type, lex,
  Lexicalize(PLACE-VIA, (first rec-node), (rest rec-node)))
return(surfstruct)

```

### 4.3 Implementation des inkrementellen Wegbeschreibungsagenten MOSES

Das in Kapitel 4.2 beschriebene funktionale Modell ist die Grundlage für die Implementation des inkrementellen Wegbeschreibungsagenten MOSES. MOSES ist mittels einer Multiagentenarchitektur realisiert, in der jedes einzelne Modul durch einen Agenten modelliert wird. Dabei sind Sequentialisierungsbeziehungen zwischen Modulen durch eine asynchrone bzw. synchrone Aktivierung der Module repräsentiert. Als Modell einer stadtähnlichen Umgebung wird ein dreidimensionales, geometrisches Modell des Campus der Universität des Saarlandes verwendet.

Die Implementation von MOSES besteht aus drei Teilsystemen. Das erste Teilsystem, MOSES-VIEW, ist die Implementation des *visuellen Objektauswahlmoduls*, sowie der Animationen, die es dem Agenten erlauben, sich in der Umgebung zu bewegen (vgl. Kapitel 4.1.3, 4.2.3). Das zweite Modul MOSES-PATH ist die Implementation des *Pfadsuchmoduls* (vgl. Kapitel 4.1.2, 4.2.2). Letztlich wird das raumbezogene Wissen im Teilsystem MOSES-SPEAK etabliert und in sprachliche Beschreibungen umgesetzt<sup>80</sup> (vgl. Kapitel 4.1.4, 4.2.4, 4.1.6, 4.2.6).

Das System MOSES ist vollständig in der objektorientierten Programmiersprache LISP (Lucid-CLOS)<sup>81</sup> implementiert worden. Die graphischen Oberflächen basieren auf dem *Com-*

<sup>80</sup>MOSES ist eine nahezu vollständige Implementation des in Kapitel 4.2 beschriebenen funktionalen Modells. Einzig die Funktionen zur Etablierung räumlicher Referenzsysteme sind erst teilweise implementiert.

<sup>81</sup>In der vorliegenden Implementation erfolgt dies auf Rechnern der HP Serie 7000.



## Funktion 25: Activate-Presentation-Buffer

*Activate-Presentation-Buffer(description)*

```
Initialize-Presentation-Buffer;
ts := Det-Time-Slot(description);
Present-Description(description, ts);
Initialize-Presentation-Buffer;
```

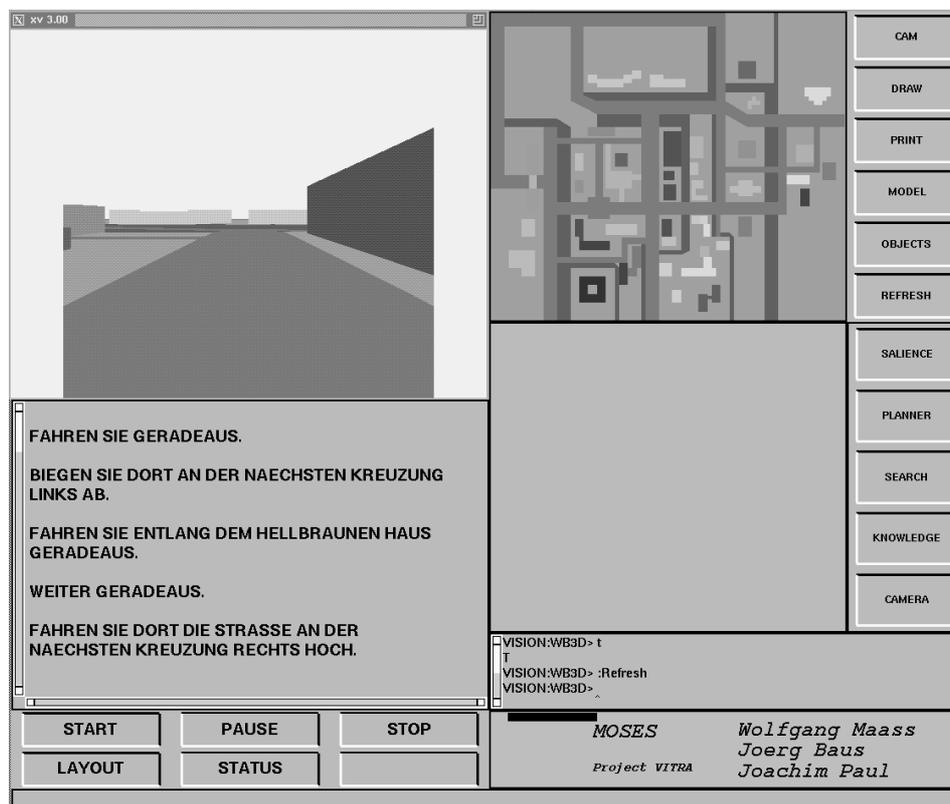


Abbildung 4.67: Bedienungsfläche von MOSES

Umgebung animiert sowie die textuelle oder verbale Beschreibung des Weges präsentiert. Zu Beginn einer Beschreibung werden vom Benutzer der Start- und Zielort per Maustastendruck auf der Straßenkarte ausgewählt. Hierzu aktiviert man zuerst die Schaltfläche *START*. Daraufhin erscheinen auf der Oberfläche die anwählbaren Knoten. Nachdem zwei Knoten ausgewählt worden sind, beginnt die Beschreibung des Weges. Neben dieser verfügt jedes Untersystem über eine eigenständige Bedienungsfläche (vgl. die nachfolgenden Abschnitte). Diese Oberflächen der Teilsysteme können aus der Gesamtoberfläche über die Schaltflächen *SALIENCE*, *KNOWLEDGE* und *SEARCH* aktiviert werden. Die anderen Schaltflächen sind für Kontrollzwecke in die Oberfläche integriert (*CAMERA*, *CAM*, *PRINT*, *OBJECTS*, *DRAW*, *MODEL*). Der Neuaufbau des aktuellen Bildes erfolgt über die Schaltfläche *REFRESH*.

MOSES basiert auf der Multiagenten-Architektur CHANNELS (vgl. [NJ94]), in der jedes Modul der funktionalen Architektur als eigenständiger Agent modelliert ist. Agenten interagieren durch Versenden von Nachrichten miteinander. Jede Nachricht ist durch die Angabe des Senders, des Empfängers sowie den Kommunikationstyp, -modus und den Inhalt bestimmt (vgl. [NJ94]). Es werden drei Kommunikationstypen unterschieden (*inform*, *ask* und *reply*). Hier-

durch erhält ein Agent Information von einem anderen Agenten, wird ein Agent durch eine Anfrage aktiviert bzw. wird eine Antwort auf eine Anfrage zurückgesendet. Entsprechend den Sequentialisierungsbedingungen ist der Modus vorgegeben, in der eine *ask* Nachricht versendet werden kann. Besteht zwischen den Modulen, welche jeweils durch einen Agenten instantiiert werden, eine sequentielle Beziehung, so sind nur synchrone *ask* Nachrichten möglich. In diesem Fall wird der sendende Agent solange in einen Wartezustand versetzt, bis auf die Anfrage eine *reply* Nachricht des empfangenden Agenten gesendet wird.

Ein Beispiel für die Interaktion zwischen Agenten ist in Abbildung 4.6 gezeigt, die einen Ausschnitt eines Protokolls zur Generierung einer inkrementellen Wegbeschreibung repräsentiert.

```

...
AGENT-2 sends SYNCH ASK to (AGENT-3) with Job-id JOB-128 and Content
(RIM)
AGENT-3 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-128 and
Content NIL
AGENT-2 sends ASYNCH ASK to (AGENT-6) with Job-id JOB-129 and
Content (WAKE-NAVIGATION (QUOTE (:SP059 :SP023 :SP024 :SP025
:SP027 :SP026)))
#<Vision #X40C60B6E>)
AGENT-2 sends SYNCH ASK to (AGENT-4) with Job-id JOB-130 and Content
(NLP)
AGENT-4 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-130 and
Content NIL
AGENT-6 sends SYNCH ASK to (AGENT-7) with Job-id JOB-131 and
Content (V3-CLIM::GET-SALIENT-OBJECT)
AGENT-5 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-139 and
Content :KRISTALLOGRAPHIE
AGENT-2 sends SYNCH ASK to (AGENT-3) with Job-id JOB-132 and Content
(RIM)
AGENT-5 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-139 and
Content :KRISTALLOGRAPHIE
AGENT-3 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-132 and
Content NIL
AGENT-2 sends SYNCH ASK to (AGENT-4) with Job-id JOB-133 and Content
(NLP)
AGENT-4 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-133 and
Content NIL
AGENT-2 sends SYNCH ASK to (AGENT-4) with Job-id JOB-134 and Content
(NLP)
AGENT-4 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-134 and
Content NIL
AGENT-2 sends SYNCH ASK to (AGENT-3) with Job-id JOB-135 and Content
(RIM)
AGENT-3 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-135 and
Content NIL
AGENT-2 sends ASYNCH ASK to (AGENT-6) with Job-id JOB-136 and
Content (WAKE-ANIMATION (QUOTE (:SP072))) #<Vision #X40C60B6E>)
AGENT-6 sends SYNCH ASK to (AGENT-7) with Job-id JOB-137 and
Content (V3-CLIM::GET-SALIENT-OBJECT)
AGENT-5 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-139 and
Content :CHEMIE
AGENT-2 sends SYNCH ASK to (AGENT-3) with Job-id JOB-138 and Content
(RIM)
AGENT-5 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-139 and
Content :CHEMIE
...

```

Tabelle 4.6: Beispiel für eine Nachrichtensequenz zwischen Agenten in MOSES

### 4.3.1 Prozesse der inkrementellen Pfadsuche

Die Implementation des Pfadmoduls basiert auf einem inkrementellen, heuristischen Suchalgorithmus, ähnlich dem RTA\*-Algorithmus von Korf (vgl. [Kor90], [RW91]). Dieses Modul bestimmt lokal optimale Pfade (vgl. Kapitel 4.2.2). Die Repräsentation des Suchraums unterteilt sich in Knoten- und Kantenrepräsentationen, welchen im einzelnen strukturspezifische Merkmale zugeordnet sind.

Die Merkmale gehen in die Bestimmung der lokalen und der geschätzten Kosten ein. Grundlegend ist eine topographieerhaltende Repräsentation der Umgebung, die durch die Zuordnung von Koordinaten erfolgt. Darüber ist ein Maß für die metrische Distanz zwischen zwei Knoten definiert. Über die Distanz ist wiederum das Konzept des Korridors definiert (vgl. Abbildung 4.68).

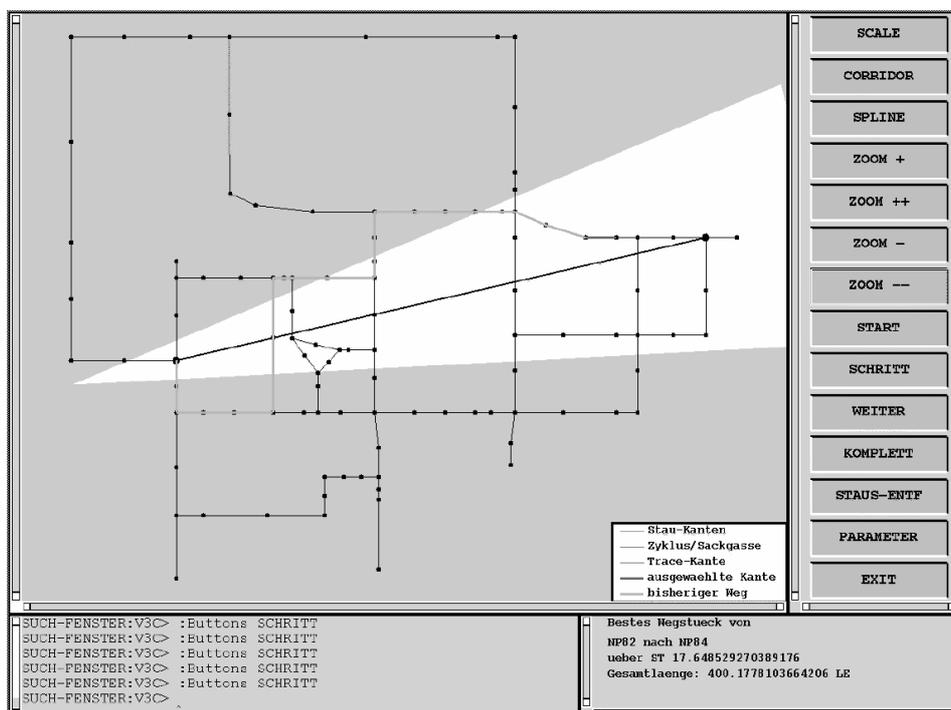


Abbildung 4.68: Oberfläche des Systems MOSES-PATH mit eingezeichnetem Korridor

Das Verhalten des inkrementellen Pfadsuchprozesses läßt sich über zahlreiche Parameter, wie den Öffnungswinkel des Korridors, die Größe der Vorausschau sowie die Gewichtung von Kantentypen modifizieren.<sup>82</sup> Alle Parameter können zur Laufzeit verändert werden, so daß eine hohe Flexibilität erreicht wird. Über Trace-Ausgaben läßt sich das Verhalten des Pfadprozesses detailliert überprüfen.

Über die Schaltfläche *SCALE* kann die Auflösung der Karte modifiziert werden. Der Korridor kann über *CORRIDOR* in seinen Parametern, wie Öffnungswinkel und das Verhältnis zwischen der Strecke vom Start- zum Zielort zur Länge des Korridors, verändert werden. In die Karte kann über vier Schaltflächen hinein- bzw. hinausgezoomt werden (*ZOOM+*, *ZOOM++*, *ZOOM-*, *ZOOM--*). Zu Beginn einer Suche wird die Schaltfläche *START* aktiviert, woraufhin der Benutzer den Start- und Zielknoten auswählen kann. Entsprechend der Parametrisierung berechnet der Suchalgorithmus den nächsten  $n$  Schritte, sobald die Schaltfläche *WEITER* aktiviert worden ist.

<sup>82</sup>Für eine detaillierte Diskussion vgl. [AK96].

Bei Auswahl der Schaltfläche *SCHRITT* wird jeweils genau ein weiteres Pfadsegment gesucht. Der gesamte Pfad kann auf einmal durch Betätigung der Schaltfläche *KOMPLETT* berechnet und angezeigt werden. Möchte man alle Blockierungen von Knoten und Kanten entfernen, kann man die Schaltfläche *STAU-ENTF* aktivieren. Möchte man eine Kante oder einen Knoten blockieren, so wählt man das Element direkt in der Karte an. Daraufhin erscheint ein Menü, aus dem man den Typ und den Namen entnehmen, sowie das Element blockieren kann.

In Abbildung 4.69 ist ein Pfad eingezeichnet, der im Startknoten A beginnt und im Zielknoten Z endet. Bedingt durch zwei blockierte Pfadabschnitte (B1 und B2) ist der Zielknoten nicht direkt erreichbar. Aus diesem Grund muß ein Pfad außerhalb des Korridors gewählt werden. Durch den Korridor versucht der Algorithmus, Pfadabschnitte in der Nähe des Korridors zu bevorzugen, weswegen der Pfadabschnitt P gewählt wird. In nächsten Schritt ergibt die heuristische Funktion, daß dies ungünstig gewesen ist und kehrt wieder über den Pfadabschnitt Q zum Ausgangsknoten zurück.

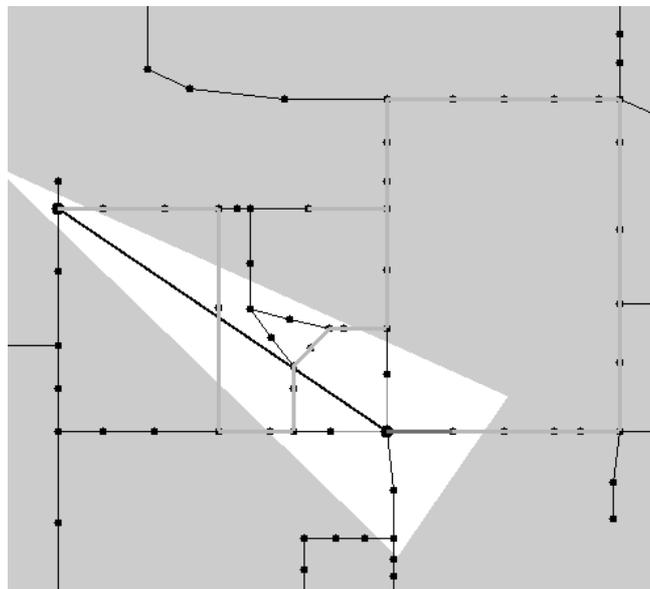


Abbildung 4.69: Darstellung eines Zielkorridors bei einer Navigation eines Fußgängers unter Berücksichtigung zweier gesperrter Pfadabschnitte *B1* und *B2*

Ein solches kurzsichtiges Verhalten wird umgangen, indem die Anzahl der im voraus bestimmten Pfadabschnitte erhöht wird. In Abbildung 4.70 ist ein Pfad mit einer Vorausschau von 1 abgebildet, in der zwei Kanten sich als ungünstig herausstellen (N/O und R/S). Bei einer Vorausschau von 1 ist der lokale Planungsaufwand minimal, führt jedoch global gesehen zu Häufungen von Zyklen und der Auswahl von ungünstigen Kanten.

Besseres Verhalten wird bei einer höheren Vorausschau erreicht. Bei identischem Start- und Zielknoten sowie einer Vorausschau von 10 Kanten wird ein wesentlich effizienterer Pfad bestimmt (vgl. Abbildung 4.71). Ein solches global gesehen besseres Verhalten, wird auf Kosten eines höheren lokalen Berechnungsaufwandes erkaufte. Setzt man die Vorausschau auf einen beliebig hohen Wert, so nähert sich das Verhalten des Algorithmus asymptotisch an den A\*-Algorithmus an. Gleichzeitig erhöht sich der lokale Berechnungsaufwand proportional zur Größe der Vorausschau. Dadurch entsteht ein "trade-off" zwischen lokaler Berechnungsökonomie und globaler Optimalität.

Unter der Annahme, daß kognitive Prozesse eine globale Optimalität zugunsten lokaler Ökonomie vernachlässigen, ist die Verwendung des inkrementellen Ansatzes unter Verwendung einer

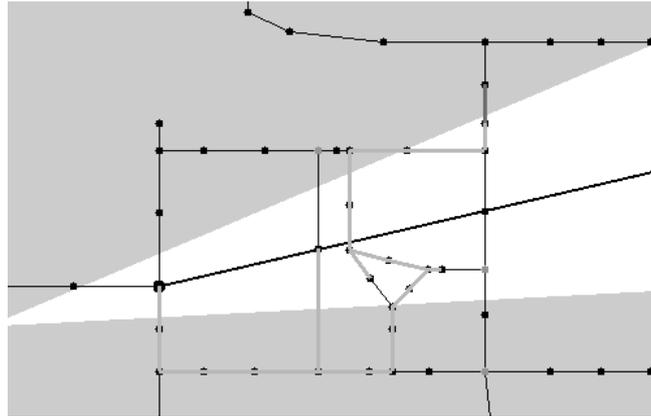


Abbildung 4.70: Beispiel einer Pfadsuche mit einer Vorausschau von einem Pfadabschnitt

begrenzten Vorausschau vorzuziehen. Die Inkrementalität der Berechnung hält die Komplexität einer Bestimmung eines Pfadabschnitts konstant. Dabei verhält sich die Komplexität proportional zum Produkt aus maximalem Verzweigungsgrad und der Länge des *look-forward*. Die gesamte Komplexität eines Pfades berechnet sich aus der Summe der Einzelevaluationen.<sup>83</sup>

### 4.3.2 Prozesse zur visuellen Objektauswahl

Die Implementation des Objektauswahl-Moduls MOSES-VIEW<sup>84</sup> basiert auf den in Kapitel 4.2.3 vorgestellten Funktionen (vgl. Abbildung 4.72).

Das System MOSES-VIEW erlaubt es, eine beliebige Situation aus dem 3D-Modell auszuwählen (Schaltfläche *SELECT-SCENE*). Die Oberfläche präsentiert links oben die ausgewählte perspektivische Sicht. Rechts oben werden die Potentialfelder durch farbig markierte Ellipsen indiziert.<sup>85</sup> Die Farbe der Ellipsen markieren den visuellen Salienzwert. In diesem Teilfenster werden die visuellen Salienzen einzelner Merkmalstypen oder aber die Kombination über alle Merkmalstypen hinweg angezeigt. Unter diesem Teilfenster wird jeweils der Querschnitt durch die Potentialfelder gezeichnet. Je höher eine einzelne Kurve ist, um so größer ist der visuelle Salienzwert des assoziierten Objektes. Links unten werden die Landmarken eingezeichnet, die aus der darüberliegenden Situation ausgewählt worden sind. Das visuell salienteste Objekt wird umrahmt und mit dem Namen des Gebäudes versehen. Die Berechnung der visuellen Salienzen der Objekte einer Situation können durch verschiedene Parameter modifiziert werden. Einer der wichtigsten Parameter ist die Ausrichtung des Fokus. Dies erfolgt über die Schaltflächen *LEFT-FOCUS*, *CENTER-FOCUS*, *RIGHT-FOCUS* sowie *SELECT-FOCUS*. Durch *SELECT-FOCUS* kann der Benutzer einen eigenen Fokusbereich per Maus definieren.

Der Blickwinkel, aus dem eine perspektivische Sicht auf das geometrische Modell des Universitäts-Campus betrachtet wird, hängt von der Lokation des Agenten (*eye point*), sowie seinem Blickpunkt (*aim point*) ab. Die Funktionen für die Modellierung des geometrischen Modells basieren auf dem System GEOMETRIX (vgl. [Her96]), sowie die Funktionen zur Bestimmung von animierten perspektivischen Sichten des Agenten auf dem System GEODISPLAY (vgl. [Mül95]). Das geometrische, dreidimensionale Modell ist durch Punkte, Kanten und Polygonzüge

<sup>83</sup>Der hier vorgestellte inkrementelle, heuristische Suchalgorithmus ist in verschiedenen Domänen, wie beispielsweise auf einem Ausschnitt des deutschen Autobahnnetzes, getestet worden (vgl. [AK96]).

<sup>84</sup>Dieses Modul ist von Jörg Baus realisiert worden (vgl. [MBP95], [Bau95]).

<sup>85</sup>Die Größe der Ellipsen ist durch die Größe des Objektes vorgegeben und besitzt aus diesem Grund keine weitere Bedeutung.

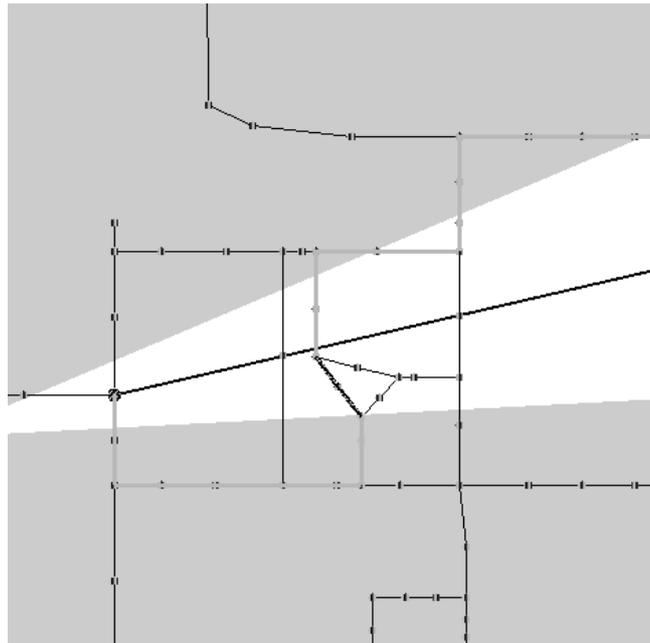


Abbildung 4.71: Beispiel einer Pfadsuche mit einer Vorausschau von zehn Pfadabschnitten

definiert, die zu komplexeren Objekten zusammengesetzt werden. In der vorliegenden Implementation sind Gebäude durch Quader approximiert. (vgl. Tabelle 4.7).

```
(defgeo :campus :point :BP0263 365.0 245.0 2.0)
(defgeo :campus :point :BP0264 375.0 245.0 2.0)
(defgeo :campus :point :BP0265 365.0 250.0 2.0)
(defgeo :campus :point :BP0266 365.0 245.0 0.0)

(defgeo :campus :point :BP0267 360.0 250.0 8.0)
(defgeo :campus :point :BP0268 380.0 250.0 8.0)
(defgeo :campus :point :BP0269 360.0 265.0 8.0)
(defgeo :campus :point :BP0270 360.0 250.0 0.0)

(defgeo :campus :point :BP0271 355.0 255.0 8.0)
(defgeo :campus :point :BP0272 360.0 255.0 8.0)
(defgeo :campus :point :BP0273 355.0 265.0 8.0)
(defgeo :campus :point :BP0274 355.0 255.0 0.0)

(defgeo :campus :point :BP0275 380.0 255.0 8.0)
(defgeo :campus :point :BP0276 385.0 255.0 8.0)
(defgeo :campus :point :BP0277 380.0 265.0 8.0)
(defgeo :campus :point :BP0278 380.0 255.0 0.0)

(defgeo :campus :box :BB0068 :WHEAT :BP0263 :BP0264 :BP0265 :BP0266)
(defgeo :campus :box :BB0069 :WHEAT :BP0267 :BP0268 :BP0269 :BP0270)
(defgeo :campus :box :BB0070 :WHEAT :BP0271 :BP0272 :BP0273 :BP0274)
(defgeo :campus :box :BB0071 :WHEAT :BP0275 :BP0276 :BP0277 :BP0278)

(defgeo :campus :object :INFORMATIK '(:BB0068 :BB0069 :BB0070 :BB0071))
```

Tabelle 4.7: Geometrische Definition des Objekts „Informatik“

In der vorliegenden Implementation von MOSES ist das zugrundeliegende Modell eine geometrische Repräsentation des Campus der Universität des Saarlandes. Die Grundfläche besteht aus einer Menge von planaren Polygonzügen. Weitere Elemente sind Pfade, Entscheidungspunkte

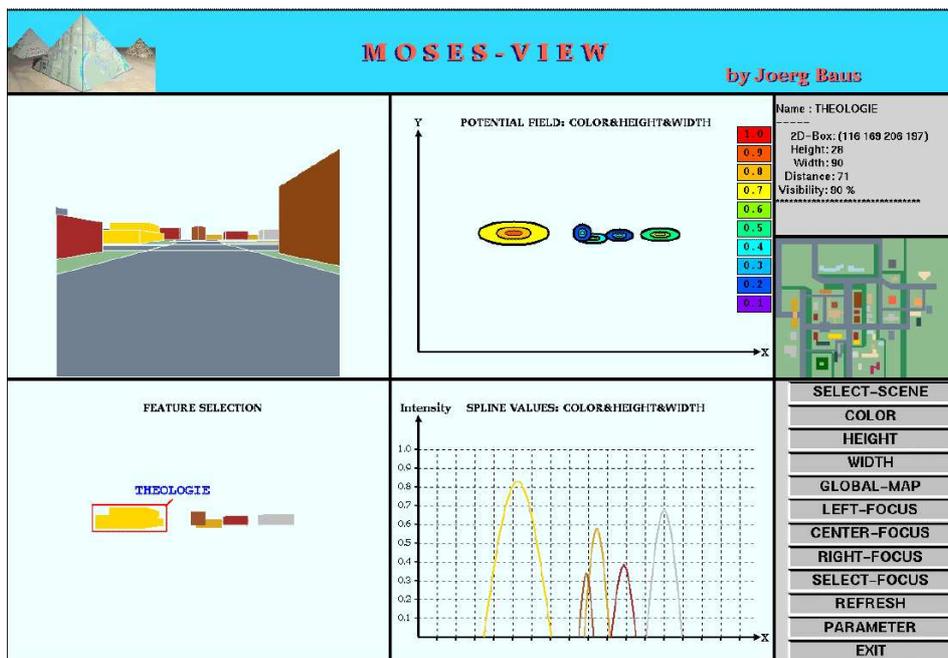


Abbildung 4.72: Bedienungs Oberfläche des Subsystems MOSES-VIEW

(Kreuzungen, Plätze und Kreisverkehre) und Landmarken. Basierend auf dem Konzept eines dreidimensionalen Objektes sind dem Konzept einer *Landmarke* (landmark) weitere Eigenschaften wie ein Name, eine Farbe sowie eine Koordinate des Schwerpunkts zugeordnet (vgl. Tabelle 5).

```
(defconcept :campus :landmark
  :name :INFORMATIK
  :color-name :hellbraun
  :center (gx2::make-3d-point :x 370.0 :y 255.0 :z 0.0))
```

Beispiel 5: Beispiel einer Instanz einer Landmarke

Ein *Pfadabschnitt* (street-polygon) ist durch ein Rechteck definiert, für das vier 3D-Koordinaten angegeben werden (vgl. die Datenstruktur in Tabelle 6). An jeden Pfadabschnitt grenzen mindestens ein, meist jedoch zwei Pfadabschnitte an (vgl. Tabelle 6).

```
(defconcept :campus :street-polygon :SP001
  :points '(:GR0163 :GR0007 :GR0002 :GR0173)
  :prev ':SP000
  :next ':SP049)
```

Beispiel 6: Beispiel einer Instanz eines Pfadabschnitts

*Entscheidungspunkte* (street-crossing) werden wie Pfadabschnitte als Rechteck definiert. An einen Entscheidungspunkt können jedoch eine beliebige Anzahl von Pfadabschnitten angrenzen (seg-1,..., seg-n) (vgl. Tabelle 7).

Grundlage für die Berechnung visueller Salienzen von Objekten ist die Bestimmung der Sichtbarkeit einzelner Objekte. Objektbezogene Flächen bezüglich des aktuellen *eye point* und *aim*

```
(defconcept :campus :street-crossing :SP027
  :points '(:GR0128 :GR0123 :GR0124 :GR0129)
  :seg-1 ' :SP031
  :seg-2 ' :SP028
  :seg-3 ' :SP025
  :seg-4 ' :SP026)
```

### Beispiel 7: Beispiel einer Instanz eines Entscheidungspunktes

*point* werden über eine interne Repräsentationsstruktur (BSP-Baum) entsprechend ihrer Distanz geordnet. Aus operationaler Sicht werden Flächen von Objekten, die weiter entfernt sind, vor solchen gezeichnet, die näher sind. Dadurch wird erreicht, daß Überlappungen korrekt gezeichnet werden (vgl. [Bau95]). Das Ergebnis ist zwar eine korrekte Darstellung der Situation bezüglich *eye point* und *aim point*, jedoch fehlt eine explizite Repräsentation der letztendlich sichtbaren Anteile der gezeichneten Flächen. Um visuelle Salienzen berechnen zu können, ist es notwendig, die Gebiete einer Fläche zu bestimmen, die von einem Blickpunkt aus sichtbar sind.

Die Bestimmung solcher Gebiete basiert auf einer Diskretisierung der Projektionsebene, so wie sie in Kapitel 4.2.3.5 vorgestellt wird. Parallel zum Zeichnen eines Objektes werden auf einer Rasterrepräsentation durch eine homomorphe Abbildung solche Zellen mit dem Namen des zu zeichnenden Objektes gefüllt, die dem Gebiet des minimal-umschreibenden Rechtecks des Objektes auf der Projektionsfläche entsprechen. Wie in Kapitel 4.2.3.5 diskutiert, hängt der Fehler dieser Approximation von der Größe einer einzelnen Zelle des Rasters ab. Flächen, deren Ausdehnung in einer Dimension kleiner als die Ausdehnung einer Zelle sind, verschwinden im ungünstigsten Falle auf der Rasterrepräsentation. Der Berechnungsaufwand einer Rasterrepräsentation steigt umgekehrt-proportional zur Größe einer Zelle.<sup>86</sup> Mittels der Rasterrepräsentation werden sichtbare lineare Merkmale eines Objektes bestimmt (vgl. Abbildung 4.21 in Kapitel 4.2.3).<sup>87</sup> Die Rasterrepräsentation ist dadurch eine explizite Repräsentation einer Situation.

In sequentieller Weise<sup>88</sup> werden situationsabhängige Mittelwerte rekursiver und linearer Merkmale etabliert (vgl. Kapitel 4.2.3.6, 4.2.3.7). Bedingt durch die Effizienz der Verarbeitung wird in der Implementation ein rechteckiger *räumlicher Fokus* verwendet.<sup>89</sup> Der räumliche Fokus definiert den Ausschnitt einer Situation, in dem visuelle Salienzen bezüglich Objekten berechnet werden. Besteht eine Intention an einem Entscheidungspunkt links abzubiegen, wird der räumliche Fokus relativ zur Frontalrichtung des primären Referenzsystems nach links verschoben.<sup>90</sup> Insgesamt sind in MOSES drei verschiedene Fokussierungsmöglichkeiten implementiert. Der Agent kann entweder nach links, rechts oder geradeaus fokussieren (vgl. [Bau95]).

Für jeden Merkmalstyp wird, entsprechend der “feature integration theory”, eine eigenständige Rasterrepräsentation (Merkmalskarte) etabliert (vgl. Kapitel 2, 4.2.3). Bezüglich einer solchen Merkmalskarte werden in einem räumlichen Fokus, entsprechend den in Kapitel 4.2.3 vorge-

<sup>86</sup>Eine mögliche Adaption von MOSES ist es, die Größe der Zellen ressourcenabhängig festzulegen. Je weniger Zeit zur Verfügung steht, um so geringer ist die Auflösung auf der Rasterebene. Dies hat Auswirkungen auf die Berechnung visueller Salienzen und dadurch auf die Genauigkeit der Analyse der gesamten Situation. Steht mehr Zeit zur Verfügung kann die Situation besser, d.h. mit höherer Auflösung, analysiert werden.

<sup>87</sup>Da in der vorliegenden Implementation von einer homogenen Verteilung rekursiver Merkmale bezüglich eines Objektes ausgegangen wird, sind diese auf Rasterebene und müssen nicht mehr neu bestimmt werden.

<sup>88</sup>Einer nebenläufigen Implementation steht kein prinzipielles Hindernis im Wege. Da aber die Implementation von MOSES-VIEW für die Evaluation auf einem herkömmlichen sequentiellen Rechner geschrieben ist, ist die Sequentialität inhärent.

<sup>89</sup>Andere Formen der Foki, wie z.B. kreisförmige, sind denkbar, erfordern jedoch bezüglich einer Rasterrepräsentation einen erhöhten Berechnungsaufwand.

<sup>90</sup>In der aktuellen Implementation ist die Ausrichtung des räumlichen Fokus statisch vorgegeben. Flexible, d.h. situationsabhängige Festlegungen, von räumlichen Foki sind in einer späteren Implementation vorgesehen.

stellten Funktionen, visuelle Salienzen bezüglich rekursiver Merkmale etabliert. Bezüglich dem rekursiven Merkmal *Farbe* werden visuelle Salienzen, entsprechend der an die MacAdamschem Ellipsen angepaßten, durch Abstände von Farbvektoren entsprechend dem von MacAdam vorgeschlagenen CIE-UCS-Diagramms (uniform chromaticity scale) (vgl. [Mac37], [WS82], [JW75]).<sup>91</sup>

Für Farbmerkmale lassen sich relative Distanzen zu dem situationsbezogenen Mittelwert ableiten, die den relativen visuellen Salienzwert bezüglich dieses Merkmalstyps angeben. Die visuellen Salienzen rekursiver Merkmale werden für jedes Objekt verglichen. Hierzu werden alle visuellen Salienzen von Objekten im Fokus in ein normiertes Koordinatensystem eingetragen. Die Höhe der visuellen Salienzkurve repräsentiert den relativen visuellen Salienzgrad. Hierbei werden Schwellwerte evaluiert, die Objekte mit zu geringen visuellen Salienzen entfernen.<sup>92</sup> Bedingt durch die *lokale* und *globale Verdichtung* von Farbmerkmalen ist die Auffälligkeit des Merkmals *Farbe* insgesamt für die Berechnung der visuellen Salienz eines Objektes von der Verteilung der Farbmerkmalswerte abhängig (vgl. Kapitel 4.2.3.6). Zur Berücksichtigung der globalen Verteilung der Farbmerkmalswerte werden visuelle Salienzen an die globale Verdichtung angepaßt, woraus unterschiedliche visuelle Salienzen eines Farbmerkmals resultieren.

Für die Berechnung visueller Salienzen der lineareren Merkmale *Höhe* und *Breite* wird von der geometrischen Ausdehnung der sichtbaren Flächen auf der Projektionsebene ausgegangen. Hierzu werden Mittelwerte bestimmt und mit den Merkmalswerten verglichen. Hieraus lassen sich visuelle Salienzen linearer Merkmale ableiten (vgl. Kapitel 4.2.3.7). Bei der Berechnung von visuellen Salienzen linearer Merkmale werden domänenabhängige Schwellwerte dazu verwendet, Objekte mit einem zu geringen absoluten Merkmalswert nicht zu berücksichtigen.<sup>93</sup> Letztlich werden typspezifische visuelle Salienzen über die Rasterrepräsentation integriert. Entsprechend den Gewichtungen der Merkmalstypen werden die visuellen Salienzen eines Objektes etabliert. Über die visuellen Salienzen der Objekte im räumlichen Fokus einer Situation wird eine Ordnung der Objekte etabliert und an den Raumintegrationsagenten (RI-M) weitergeleitet.

Das Modul MOSES-VIEW evaluiert bezüglich einer beliebigen Perspektive visuell sichtbare Objekte, die Größe der sichtbaren Flächen, objektbezogene visuelle Salienzen rekursiver und linearer Merkmale sowie deren Integration (vgl. Kapitel 4.2.3). Das in MOSES integrierte Objektauswahl-Modul ist auch als eigenständiges System über eine graphische Benutzungsoberfläche ansteuerbar (vgl. Abbildung 4.72). Über diese Oberfläche können beliebige Sichten ausgewählt und sichtbare Objekte bezüglich ihrer visuellen Salienzen geordnet werden (vgl. [Bau95]). Die Oberfläche erlaubt, einen direkten Vergleich zwischen einer perspektivischen Sicht und merkmalspezifischen und integrierten visuellen Salienzen vorzunehmen.

---

<sup>91</sup>In der aktuellen Version basiert die Evaluation von farbbezogenen visuellen Salienzen auf einer RGB-Zerlegung. Dies ist solange sinnvoll, wie von einem konstanten Sättigungsgrad und konstanter Helligkeit der Farben ausgegangen werden kann. Eine Berücksichtigung dieser Parameter liegt außerhalb des Rahmens dieser Arbeit, da es vorrangig um eine prinzipielle Berücksichtigung von Farbmerkmalen geht (hierzu vgl. [Gap96]). Eine Verwendung von Farbzerlegungen nach alternativen Systemen kann in die in Kapitel 4.2.3.6 vorgestellten Funktionen in einfacher Weise erfolgen. In weiteren Arbeiten soll untersucht werden, ob psychologisch motivierte Farberscheinungsmodelle eine bessere Übereinstimmung mit empirischen Farbwahrnehmungstheorien liefern. Beispiele für solche Modelle sind das Munsell Farbsystem (in [WS82]), das Natural Color System (NCS) (vgl. [Här66]) und das DIN Farbsystem (vgl. [DIN80]).

<sup>92</sup>In der aktuellen Version sind diese Schwellwerte statisch. Wiederum in Bezug auf ressourcenbeschränkte Umgebungen sollen in einer Erweiterung dieses Ansatzes die Schwellwerte flexibel anpaßbar sein.

<sup>93</sup>Dies steht im Gegensatz zur "feature integration theory", jedoch ist es nicht sinnvoll davon auszugehen, daß ein kleiner Ball aufgrund seiner stark vom Mittelwert der Objekte einer Situation abweichenden Breite und Höhe besonders auffällig sein soll.

### 4.3.3 Prozesse zur Verarbeitung räumlichen Wissens

Die Implementation der Funktionen des Raumintegrationsmoduls ist Teil des Subsystems MOSES-SPEAK.<sup>94</sup> In MOSES-SPEAK wird Routen- und Landmarkenwissen über eine Situation aufgebaut. Räumliches Wissen basiert wesentlich auf dem Konzept eines Referenzsystems, welches Objekte, wie Landmarken und Pfadelemente, integriert (vgl. Abbildung 4.73).

Im deiktischen Fall werden räumliche Relationen bezüglich eines primären oder sekundären Referenzsystems etabliert. Intrinsische Relationen zwischen Objekten basieren auf der Verwendung eines virtuellen Referenzsystems. Da für die Etablierung räumlicher Relationen Funktionen verwendet werden, die grundsätzlich von einem im Referenzobjekt verankerten Referenzsystem ausgehen (vgl. Kapitel 4.2.4.1 und [Her96], [Gap94]), wird aus Implementationsgründen bei der Etablierung der räumlichen Relationen #in-front-of# und #behind# ein extrinsisches Referenzsystem und für die Relationen #left-of# und #right-of# ein um 180° gedrehtes Referenzsystem angenommen. Eine solche Verwendungsweise von virtuellen Referenzsystemen bildet die Funktionen nach, die in Kapitel 4.2.4.1 vorgestellt worden sind.

Räumliche Relationen zwischen Objekten werden als Listen verwaltet, die aus dem Namen, den Referenzen auf die Objekte und dem Anwenbarkeitsgrad der Relation bestehen.

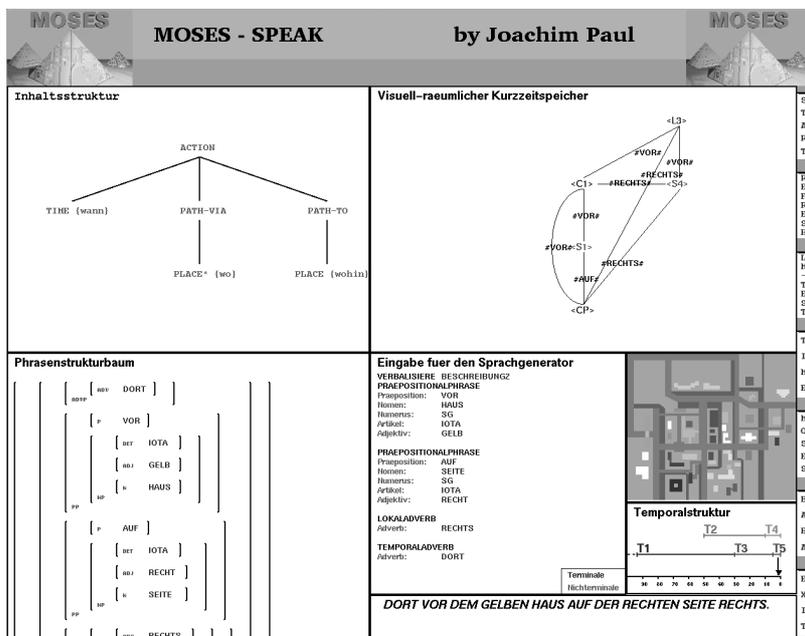


Abbildung 4.73: Die Bedienungs Oberfläche des Systems MOSES-SPEAK

Das Subsystem MOSES-SPEAK ist sowohl in MOSES als Prozeß über einen Agenten integriert, als auch als eigenständiges System verwendbar (vgl. Abbildung 4.73). In MOSES-SPEAK erfolgt der dynamische Aufbau räumlichen Wissens, der Auswahl von Inhaltsstrukturen und deren Umsetzung in Phrasenstrukturen. Weiterhin wird der geschätzte Zeitrahmen  $T_{guess}$  angezeigt. Das Verhalten von MOSES-SPEAK ist über Parametereinstellungen modifizierbar (vgl. [Pau95]).

Ebenso wie die Bedienungs Oberfläche von MOSES verfügt MOSES-VIEW über die Möglichkeit eine inkrementelle Wegbeschreibung durch Aktivierung der Schaltfläche START zu starten. Daraufhin erscheint in der kleinen Karte rechts unten das Netz, in dem der Benutzer den Start- und Zielknoten auswählen kann. Auf der Oberfläche wird links oben jeweils die aktuell ausgewählte

<sup>94</sup>Das Subsystem MOSES-SPEAK ist von Joachim Paul implementiert worden ([Pau95]).

Inhaltsstruktur angezeigt. Rechts oben befindet sich die aktuelle räumliche Konfigurationsbeschreibung des visuell-räumlichen Kurzzeitspeichers. Links unten wird in Klammernotation die Phrasenstruktur der aktuellen Beschreibung gezeigt sowie rechts daneben die Eingabe für die Komponente zur Bestimmung einer Terminalkette. Unter der Karte wird der aktuelle Zeitrahmen durch einen Pfeil indiziert. Rechts unten wird die aktuell gespeicherte Beschreibung angegeben. Diese wird ausgegeben, sobald der Agent die Grenze des aktuellen Zeitrahmens erreicht hat.

#### 4.3.4 Prozesse zur Generierung natürlicher Sprache

Der zweite Bestandteil des Subsystems MOSES-SPEAK dient der Generierung sprachlicher inkrementeller Wegbeschreibungen. Entsprechend des Zeitrahmens  $T_{\text{guess}}$  werden geeignete Inhaltsstrukturen ausgewählt und instantiiert. Hierzu wird zeitliches, intentionales und räumliches Wissen verwendet, um Inhaltsfragmente zu ersetzen. Zwischen einem Inhaltsfragmenttyp und entsprechenden Wissenstrukturen sind Abhängigkeiten definiert, die vorgeben, durch welche Wissensstruktur ein Inhaltsfragment ersetzt werden kann. Eine WO-K Kategorie wird ausschließlich durch eine Beschreibung einer anwendbaren Landmarke oder einem Entscheidungspunkt ersetzt. WOHIN-K und WAS-K Kategorien auf pfadbezogenen Intentionen, wohingegen zeitliches Wissen die WANN-K Kategorie instantiiert.

Die Umsetzung von Wissensstrukturen in Oberflächenstrukturen erfolgt regelbasiert auf der Grundlage eines funktional-grammatikalischen Ansatzes. Das Ergebnis ist eine Oberflächenstruktur der Beschreibung. In MOSES erfolgt die Umsetzung in eine sprachliche Beschreibung optional über eine Anbindung an HPSG, durch Verwendung eines TAG-Generators (vgl. [FS92]) oder einen einfachen Sprachgenerator. Die Ausgabe der inkrementellen Wegbeschreibung erfolgt sowohl textuell auf der Oberfläche, als auch optional verbalsprachlich.

#### 4.3.5 Beispiel einer automatisch generierten inkrementellen Wegbeschreibung

In Abbildung 4.74 ist eine Beispielsequenz abgebildet. Für dieses Beispiel ist MOSES als Fußgänger konfiguriert worden, der über gute sprachliche Fähigkeiten verfügt.<sup>95</sup> Weiterhin integriert MOSES nach Möglichkeit Landmarken in seinen Beschreibungen.

Auf der linken Seite ist die Lokation von MOSES durch einen gelben Punkt auf der Karte des Campusgeländes eingezeichnet. MOSES ist auf seinem Weg vom Informatikgebäude zur Universitätsbibliothek. In Abbildung 4.74 sind die Lokationen von drei Beschreibungen durch eine Graphik auf der linken Seite illustriert. Zuerst wird eine Geradeausbewegung "über eine Kreuzung hinweg" beschrieben. Entsprechend den Ergebnissen der empirischen Untersuchung wird darin keine weitere Landmarke referenziert. In der zweiten Beschreibung wird ein Linksabbiegevorgang indiziert. Die Lokation der Aktion ist durch die Referenzierung einer Landmarke (das dunkelbraune Haus) genauer spezifiziert. Unterstützend erhält der Rezipient nachfolgend die Information, daß er sich immer noch auf dem richtigen Pfad befindet und weiter geradeaus bewegen soll. Um dies zu unterstützen wird eine Landmarke (das hellbraune Haus) referenziert. Kurz vor der Ausführung der Aktion wird die Aktion noch einmal kurz angedeutet. Da dieser Pfadabschnitt verhältnismäßig lang ist, wird nochmals bestätigt, daß man den richtigen Weg verfolgt. Vor dem Erreichen der nächsten Kreuzung wird die Lokation der auszuführenden Aktion (rechts abbiegen) durch einen deiktischen Ausdruck (dort) spezifiziert. Daß zwischen dem aktuellen Pfadabschnitt und dem nachfolgenden ein Höhengefälle besteht, wird durch die adverbiale

<sup>95</sup>Gute Beschreibungen unterscheiden sich von normalen darin, daß sie grammatikalisch korrekt sind. Die Mehrzahl der Probanden generierten jedoch Beschreibungen, die teilweise nur unzureichend die Grammatik der deutschen Sprache berücksichtigten (vgl. das einleitende Beispiel in Abschnitt 1.1). Implementatorische Aspekte hierzu sind in [Pau95] genauer beschrieben.

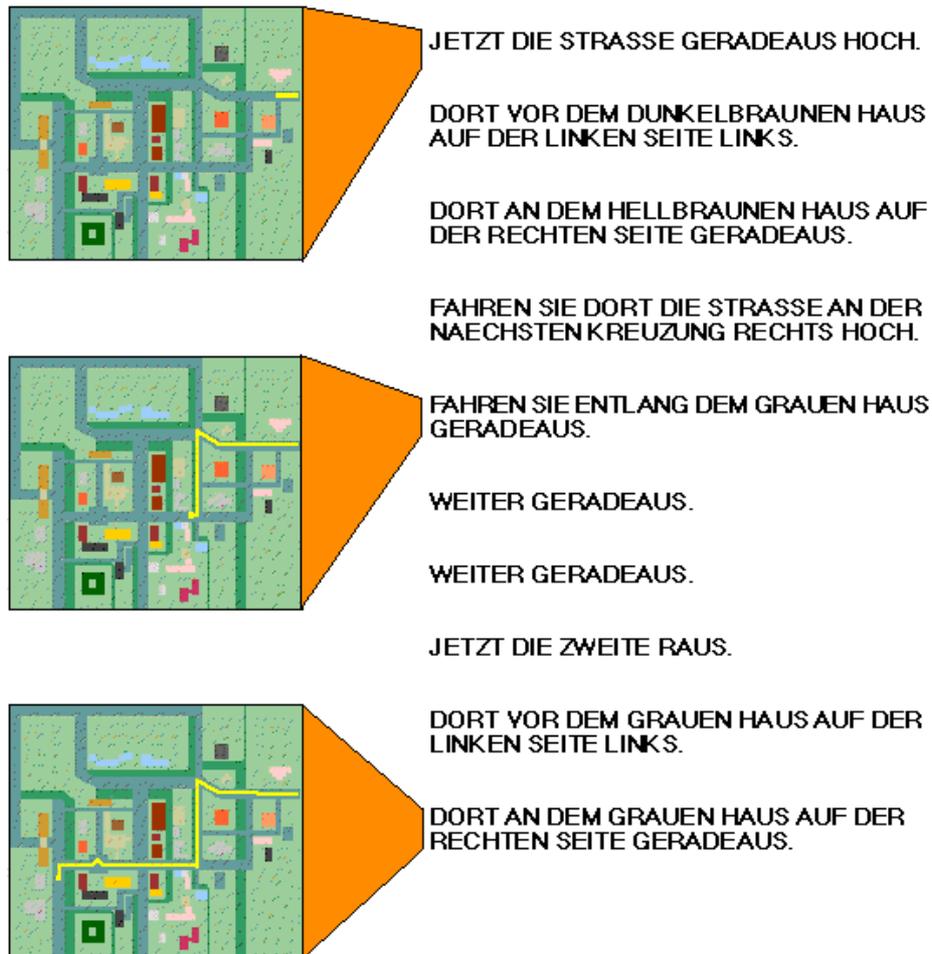


Abbildung 4.74: Sequenz einer von MOSES generierten inkrementellen Wegbeschreibung

Ergänzung „hoch“ angegeben. Danach werden drei unterstützende Beschreibungen gegeben, bevor der Agent einen Kreisverkehr erreicht. In diesem muß der Agent die zweite Ausfahrt nehmen, was entsprechend sprachlich beschrieben wird. Nachfolgend ist eine Linksabbiegeaktion auszuführen, welcher MOSES durch Nennung einer Landmarke auf der linken Seite beschreibt. Bei erhöhter Geschwindigkeit werden die Beschreibungen kürzer. Dies ist zum einen darauf zurückzuführen, daß aus Zeitgründen weniger Landmarken selektiert und integriert werden können, zum anderen auf die Verwendung unterschiedlicher Inhaltsstrukturen. Vergleicht man die Beschreibungen aus Beispiel 4.74 mit denen aus Beispiel 4.74 fällt auf, daß die Anzahl der Beschreibungen im zweiten Fall geringer ist. Dies rührt daher, daß weniger unterstützende Beschreibungen gegeben werden müssen. Hieran ist erkennbar, daß im zweiten Beispiel mehr minimale Beschreibungen verwendet werden, da jeweils der zeitliche Rahmen nicht ausreichend ist, um erweiterte Beschreibungen zu generieren.

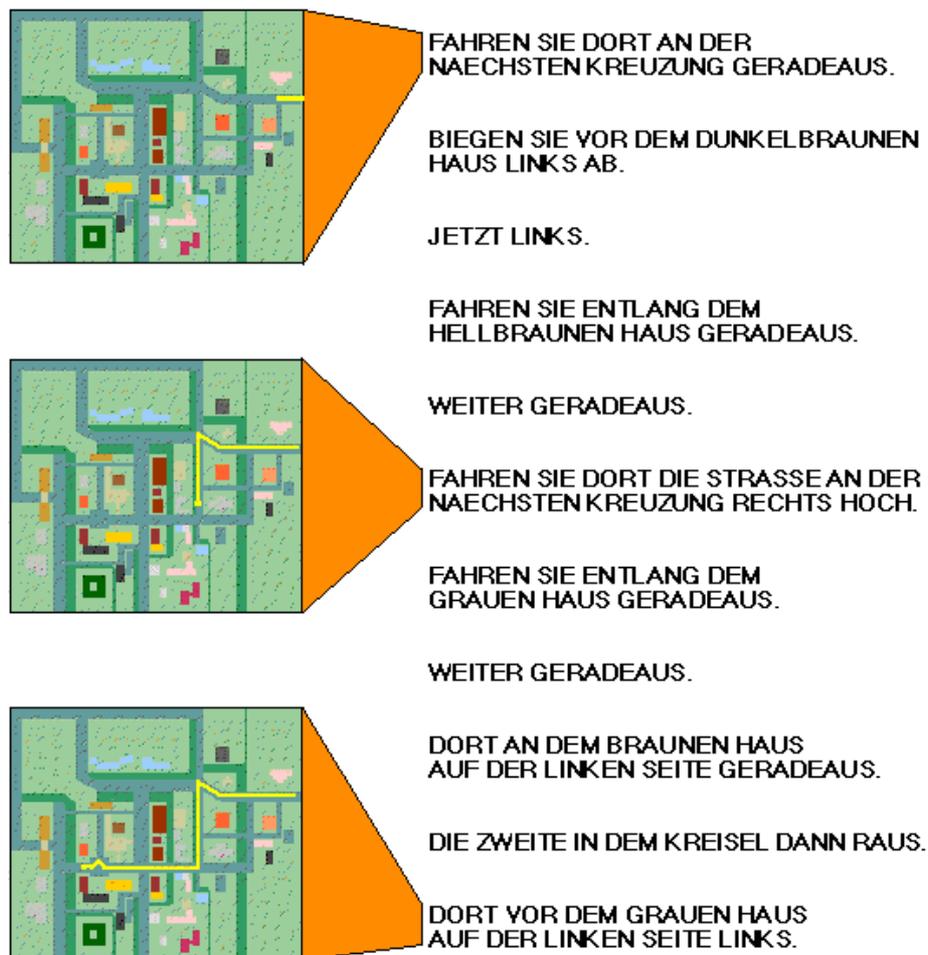


Abbildung 4.75: Sequenz einer von MOSES generierten inkrementellen Wegbeschreibung bei erhöhter Geschwindigkeit



## Kapitel 5

# Erzielte Ergebnisse, Technologische Betrachtung und Erweiterungen

In der vorliegenden Arbeit wurde ein agentenbasierter Ansatz zur Generierung inkrementeller Wegbeschreibungen in dynamischen 3D-Umgebungen vorgestellt. Damit wurde von Seiten der Grundlagenforschung ein Beitrag zur Integration von Leistungen der visuellen Datenverarbeitung, der Verarbeitung räumlichen Wissens und der Generierung natürlicher Sprache unter Einnahme einer kognitionswissenschaftlichen Perspektive geleistet. In diesem Kapitel werden die zentralen Ergebnisse dieser Arbeit zusammengefaßt. Anschließend wird der technologische Kontext im Hinblick auf Navigationssysteme vorgestellt und bezüglich dieser der praktische Nutzen des Modells bezüglich realer und virtueller Umgebungen skizziert. Abschließend werden Aspekte angeführt, in welcher Weise Erweiterungen des vorgestellten formalen Modells möglich sind.

### 5.1 Wissenschaftlicher Beitrag dieser Arbeit

Eine verbale Beschreibung dessen, was via visueller Perzeption wahrgenommen wird, ist eines der komplexesten aber gleichzeitig auch interessantesten Fragestellungen aus Sicht der Künstlichen Intelligenz und anderer kognitionswissenschaftlich orientierter Forschungsrichtungen. Da dieser gesamte Problembereich momentan als viel zu komplex erscheint, um als Ganzes betrachtet werden zu können, erscheint es als einzige sinnvolle Herangehensweise, Teilprobleme zu betrachten. Eingebettet in diesen Kontext, ist es erstmals für die Domäne der inkrementellen Wegbeschreibungen durch das vorliegende formale Modell gelungen, eine Vielzahl von Leistungen zu integrieren, welche es erlauben, komplexe Raumbeschreibungen bei gleichzeitiger Bewegung des Sprechers zu generieren. Zusätzlich wurde anhand der verwendeten Methodologie exemplarisch aufgezeigt, in welcher Weise empirische Untersuchungen bei der Entwicklung formaler Modelle und letztendlich von Implementationen verwendet werden können.

Folgende Punkte fassen den wissenschaftlichen Beitrag dieser Arbeit zusammen:

#### 1. Empirie - Modellbildung - Implementation

Diese Arbeit ist unter Verwendung der Methodologie *Empirie - Modellbildung - Implementation* entstanden. Diese der Kognitionswissenschaft entstammende Vorgehensweise erfordert die Berücksichtigung empirischer Befunde bei der Entwicklung eines Modells. Hierdurch lassen sich die Stärken empirischer Untersuchungen, die ihrem Wesen nach punktuell sind, mittels formaler Modelle in kohärenter Weise integrieren. Diese Methodologie ist bisher in Arbeiten im Bereich der Künstlichen Intelligenz vergleichsweise wenig beachtet worden. Dadurch stellt diese Arbeit auch in dieser Hinsicht einen ersten, kleinen

Schritt in der Verwendung dieser Methodologie dar. Besonders ist hierbei die Kombination aus Feld- und Laboruntersuchung hervorzuheben, die speziell an die zu untersuchenden Fragestellungen angepaßt worden ist. Weitere Arbeiten müssen zeigen, wie tragfähig diese Methodologie ist.

## 2. Sprachliche Referenzbildung in visuellen Daten

Der sprachliche Zugang zu visuellen Daten ist ein Forschungsgebiet, welches in der jüngsten Vergangenheit immer stärker an Aufmerksamkeit gewinnt. Durch diese Arbeit ist am Beispiel inkrementeller Wegbeschreibungen dieser Zusammenhang in dynamischen, dreidimensionalen Umgebungen auf der Basis eines geometrischen Modells genauer untersucht worden. Durch das integrierte Modell eines Agenten konnte gezeigt werden, wie sprachliche Referenzen in visuellen Daten verankert werden können. Bedingt durch die Dynamik des Agenten werden Referenzen und damit die Beschreibungen inkrementell angepaßt. Darüberhinaus ist der Agent durch Analyse visueller Merkmale einer Situation in der Lage Attribute von Objekten auf sprachlicher Ebene zu verankern.

## 3. Ressourcenadaptives Verhalten: Berücksichtigung von zeitlichen Restriktionen

Ein weiterer Schwerpunkt dieser Arbeit liegt darin zu zeigen, wie die Generierung von inkrementellen Wegbeschreibungen auf Beschränkungen angepaßt wird, die durch die Umgebung bzw. der Bewegung des Agenten in der Umgebung determiniert sind. Mit dieser Arbeit ist erstmals untersucht worden, welche Auswirkungen zeitliche Beschränkungen auf die Generierung von Wegbeschreibungen haben. Zwar ist die zugrundeliegende empirische Untersuchung mit Fußgängern durchgeführt worden, es ist jedoch zu vermuten, daß die gefundenen zeitkategorialen Effekte in entsprechender Weise auch für andere Bewegungsarten gelten. Am Beispiel der Ressource *Zeit* ist untersucht worden, wie eine externe, d.h. durch die Umgebung vorgegebene Ressource, vom Agenten internalisiert und verwendet wird. Basierend auf empirischen Untersuchungen ergaben sich fünf Zeitrahmen, zu denen die Probanden unterschiedlich strukturierte Beschreibungen gaben.

## 4. Erweiterung des Konzeptes des räumlichen Referenzsystems auf dynamische Aspekte

In der bisherigen Forschung zur Raumkognition sind vor allem statische Situationen betrachtet worden, in denen sich der Sprecher nicht bewegte. Durch die Eigenbewegung des Sprechers im Kontext inkrementeller Wegbeschreibungen bedingt, sind in dem vorgeschlagenen Modell die Konzepte erweitert worden. Dies gilt vor allem für Referenzsysteme. Es ist gezeigt worden, wie egozentrische und allozentrische Ausprägungen durch ein homogenes Konzept integriert und voneinander abgeleitet werden können. Das bezüglich dieser verschiedenen Referenzsysteme etablierte räumliche Wissen ist für einzelne Situationen definiert und wird aus dem räumlichen Kurzzeitspeicher entfernt, sobald der Agent eine neue Situation erreicht. Die Erweiterung von Referenzsystemen einerseits und die inkrementelle Etablierung räumlichen Wissens andererseits, ist Grundvoraussetzung für ein adaptives und effizientes Verhalten des Agenten. In dieser Arbeit wurde vor allem untersucht, wie sich das sprachliche Verhalten an die Beschränkungen, die durch die Dynamik der Situationen entstehen, adaptieren kann.

## 5. Intention geleitete Auswahl von Referenzobjekten mittels visueller Salienz

Eine der zentralen Fragen im Bereich der Verbindung visueller Datenverarbeitung und Sprachgenerierung ist, wie Referenzen zwischen Repräsentanten auf der Sprachebene mit Repräsentanda auf der visuellen Ebene verknüpft werden können. Im Kontext von Raumbeschreibungen liegt der Fokus der Untersuchungen vor allem auf der Auswahl von geeigneten Referenzobjekten. Dies ist in bisherigen Arbeiten meist unter der Prämisse einer

neutralen Betrachterposition analysiert worden. Wie es sich im Kontext inkrementeller Wegbeschreibungen zeigt, muß eine solche Position aufgegeben werden. Der Prozeß zur Auswahl von Referenzsystemen basiert einerseits auf visuellen Merkmalen, aber auch auf pfadbasierten Intentionen des Sprechers. Es ist in dieser Arbeit, aufbauend auf einer Dekompositionshypothese in visuelle Merkmale, ein Ansatz vorgestellt worden, mit dem die visuelle Salienz eines Objektes quantifiziert werden kann. Die zusätzliche Berücksichtigung von Intentionen indiziert, daß inkrementelle Wegbeschreibungen in einem kommunikativen Kontext gegeben werden, in dem sowohl der Sprecher, als auch der Rezipient Ziele verfolgen. Pfadbasierte Intentionen werden in diesem Modell dazu verwendet, einen räumlichen Aufmerksamkeitsbereich zu etablieren, in dem Objekte fokussiert werden. Dieser Aspekt ist um so wichtiger, da der Agent sich durch eine dreidimensionale Umgebung bewegt und dabei fortwährend seine Perspektive ändert.

Die Notwendigkeit zur Adaption an eine Umgebung bedingt die Verwendung von Prozessen und Repräsentationen, die eine hohe Reaktivität ermöglichen. Dies verlangt auch in zeitkritischen Situationen nach Beschreibungen, selbst wenn diese suboptimal sind. Ein solches *anytime*-Verhalten ist in die Steuerungs- und Verarbeitungsmechanismen des beschriebenen Modells explizit eingegangen.

#### 6. Inhaltskategorien inkrementeller Wegbeschreibungen

Durch Analyse des empirisch erhobenen Korpus inkrementeller Wegbeschreibungen ließen sich fünf Kategorien extrahieren, die in charakteristischer Weise zu *Inhaltsstrukturen* zusammengesetzt werden. Inhaltsstrukturen werden dynamisch in Abhängigkeit von der zur Verfügung stehenden Zeit verwendet. Entsprechend den damit verbundenen relativen Häufigkeiten dienen Inhaltsstrukturen als semantische Repräsentation für die Sprachgenerierung inkrementeller Wegbeschreibungen.

Zusammenfassend leistet das vorgeschlagene formale Modell einen Beitrag im Hinblick auf die Forderungen nach Adaptivität, Flexibilität, Effizienz und Effektivität am Beispiel inkrementeller Wegbeschreibungen.

## 5.2 Einordnung dieser Arbeit in den technologischen Kontext

In den ersten Jahren bestimmten vor allem technische Lösungen die Produkte des Softwaremarktes. Der Benutzer hatte sich auf die Bedienung der Programme einzustellen und nicht die Programme auf den Benutzer. Hierfür ist die Entwicklung von Betriebssystemen ein Beispiel. In der Anfangszeit wurden Betriebssysteme, wie beispielsweise FMS (Fortran Monitor System) oder IBSYS Betriebssystem für die 7094, verwendet, welche ausschließlich von speziell geschulten Experten bedient werden konnten. Mit dem Aufkommen von PCs erkannte man die Notwendigkeit, auch für Laien leicht zu bedienende Betriebssysteme zu entwickeln. Dies wurde in besonderer Weise von der Firma *Apple* Anfang der 80er Jahre erkannt und umgesetzt. Der dadurch ausgelöste Trend hin zu bedienerfreundlichen Betriebssystemen bestimmt seither nicht nur den PC-sondern auch den UNIX-Markt und vermehrt auch den Großrechnermarkt.

Die Folge dieser Entwicklung ist die Forderung, die in gleicher Weise für Programmiersprachen, Tabellenkalkulations- und Textverarbeitungsprogramme gilt, daß Programmsysteme sich heutzutage an die funktionalen, aber auch ergonomischen Bedürfnisse des Benutzers adaptieren sollen. Die Bedürfnisse der Benutzer lassen sich in zwei Klassen einteilen. Erstens gibt es die Klasse der Bedürfnisse, die von einer hinreichend großen Menge von Personen geteilt werden. Beispielsweise zeigt es sich, daß Benutzer auf Betriebssystemebene bevorzugt mit graphischen Elementen

umgehen. Die zweite Klasse von Bedürfnissen ist individualistisch und unterscheidet sich teilweise erheblich bei einzelnen Benutzern. Ein Beispiel hierzu sind Präferenzen für Objekteinfärbungen auf graphischen Oberflächen. Heutige Programme versuchen nahezu ausschließlich, die erste Klasse von Bedürfnissen zu decken. Bedürfnisse der zweiten Art sind wesentlich aufwendiger und besitzen inhärente Probleme. So ist es aus Datenschutzgründen nicht möglich, zentral Daten über Einzelpersonen zu speichern, um diese bei der Konfigurierung eines Programms zu verwenden. Eine andere Möglichkeit besteht darin, Benutzerprofile auf speziellen Datenträgern, wie *SmartCards*, zu speichern, die vor der Verwendung eines Programms dieses, individuell an den Benutzer angepaßt, konfigurieren. Die Repräsentation von Benutzereigenschaften wirft wiederum eigene Fragestellungen auf, die im Zusammenhang mit *Benutzermodellen* diskutiert werden ([KW89]).

Da man heute weit davon entfernt ist, zu verstehen, in welcher Weise Programme einen Benutzer maximal unterstützen, ist man auf Approximationen angewiesen, die schrittweise verfeinert werden. Ein besseres Verständnis über die Bedürfnisse erhält man durch empirische Untersuchungen. Eine grundlegende Annahme ist dabei, daß ein Programmsystem um so besser ist, je geringer der Aufwand ist, den eine Person betreiben muß, um eine bestimmte Aufgabe zu lösen. Demzufolge sind Programme so zu konzipieren, daß sie dem Benutzer situationsabhängig die Information liefern, die das Lösen einer Aufgabe mit minimalem Aufwand seitens des Benutzers unterstützen. Dies erfordert ein Verständnis besonders der kognitiven Prozesse, aber auch der Präferenzen, die sich z.B. kulturbedingt herauskristallisiert haben. Ein Beispiel hierzu ist die Bevorzugung der vis-à-vis Situation zwischen Kommunikationspartnern (vgl. Kapitel 2.5).

Vor diesem skizzierten Hintergrund ist das vorgeschlagene Modell zur Generierung inkrementeller Wegbeschreibungen auf der Basis visuell-geometrischer Daten zu verstehen. Das Modell liefert durch Prozesse und Repräsentationen, die von empirischen Befunden approximativ abgeleitet worden sind, situationsabhängige Beschreibungen. Letztendliches Ziel ist es, entsprechend der Forderung nach Effektivität, Beschreibungen derart zu generieren und zu präsentieren, daß sie vom Rezipienten als *natürlich*, *verständlich* und *adäquat* empfunden werden. Eine Beschreibung ist dann natürlich, wenn sie zum richtigen Zeitpunkt gegeben wird und kommunikative Konventionen einhält. Sie ist verständlich, wenn der Inhalt *kohärent* zu vorhergehenden Beschreibungen ist und in sich selbst *konsistent* sind. Elemente, die durch die Beschreibung referenziert werden, müssen auch vom Rezipienten bei angemessenem Aufwand *verifizierbar* sein. Letztlich ist eine Beschreibung adäquat, wenn sie sich auf den Transport der Information beschränkt, die der gegebenen Situation angemessen ist. Kommunikationsteilnehmer unterscheiden sich dabei in ihren Präferenzen. So kann eine bestimmte Wegbeschreibung für einen Touristen adäquat sein, jedoch von einem Einheimischen als zu lang und ausführlich empfunden und deshalb abgelehnt werden (vgl. [HK91]).

In der natürlichen Kommunikationssituation werden Wegbeschreibungen bezüglich Realumgebungen gegeben. Pfade sind physische Entitäten, auf denen sich Personen fortbewegen, um von einem Ausgangsort zu einem Zielort zu gelangen. Beschreibungen solcher Pfade referenzieren auf Landmarken und geben räumliche Lagebeziehungen und Bewegungsrichtungen an. Ähnliche Konzepte finden sich bei der Navigation in virtuellen Umgebungen wieder. Dies ist nicht verwunderlich, da in virtuellen Umgebungen räumliche Strukturierungen bevorzugt dazu verwendet werden, den Benutzer adäquat zu unterstützen. Inwieweit das vorgeschlagene Modell zur Generierung inkrementeller Wegbeschreibungen in realen und in virtuellen Umgebungen verwendet werden kann, wird im folgenden diskutiert.

### 5.2.1 Navigationshilfesystem in Realumgebungen

Beschreibungen von Wegen werden im alltäglichen Leben häufig gegeben. Wege, die man häufig benutzt, kennt man „im Schlaf“. In fremden Umgebungen besitzt man ein derartiges räumliches Wissen jedoch nicht.

Zu den ersten Beobachtungen, die manche Besucher Tokios als tröstlich, andere als bestürzend erleben, zählt die Erkenntnis, daß sich niemand in dieser Stadt wirklich auszukennen scheint. Kein Taxifahrer, kein Polizist, kein Fremdenführer, erst recht kein Stadtplaner kennt mehr als seinen Stadtausschnitt, individuell geschlagene Schneisen in einem Dschungel, die sich naturgemäß im Zentrum um den unbegehbaren Kaiserpalast zu Lichtungen weiten. Keinen von ihnen beunruhigt, daß nur die größten Straßen Namen führen, daß Hausnummern historisch nach Baujahr geordnet, dem Postboten geläufig sind, aber nicht einmal den Nachbarn untereinander - kurz und paradox: daß man eine unbekannte Adresse nur finden kann, wenn man schon einmal dagewesen ist. (aus FAZ, 29. Mai 1996, Nr. 123, Seite 3)

Wie Schätzungen belegen, werden jährlich ca. 7% der Fahrten mit dem Auto dadurch verschwendet, daß der Fahrer sich verfährt oder zu lange Wege benutzt. Dies macht in den USA bis zu 4 Millionen Liter Kraftstoff aus (vgl. [Bad93]). Bezogen auf Deutschland schätzt man, daß im Jahr 2000 ca. 27 Milliarden Kilometer auf diese Art zuviel gefahren werden (vgl. [RHA92]). Hierbei ist nicht eingerechnet, wieviel Millionen Liter in unnötigen Staus verbraucht werden. Ähnliches gilt für Person, die in einer Stadt beispielsweise mit dem Zug neu ankommen und denen jegliche Orientierung fehlt. In verstärktem Maße werden Navigationshilfen von Personen benötigt, die nur über eingeschränkte perzeptuell-räumliche Fähigkeiten verfügen. Solche Szenarien und Zahlen sind dafür verantwortlich, daß man sich von computerbasierten Hilfen eine erhebliche Reduktion dieser Vergeudung von Ressourcen verspricht. Schätzungen zufolge führt der Einsatz intelligenter Fahrerinformationssysteme zu einer Reduzierung der gesamten Lenkzeiten in den USA um 3,8 Millionen Stunden auf insgesamt 5,2 Millionen Stunden, des Kraftstoffverbrauchs um 4 Millionen Liter, der Emission von Kohlenwasserstoffen um 300000 Kilogramm sowie von Kohlenmonoxiden um 4 Millionen Kilogramm (vgl. [Bad93]).

Im Bereich des Transportes und der Navigation, besonders zu militärischen Zwecken, sind Technologien entwickelt worden, die computerbasierte Lösungen von Orientierungs- und Navigationsproblemen erst ermöglichen:

1. Global positioning systems (GPS)
2. Automatische Fahrzeuglokalisierung (AVL)
3. Pfadsuchalgorithmen
4. Elektronische Karten und andere Datenmengen
5. Fahrzeugbasierte Navigations- und Fahrerinformationssysteme

In Bezug auf die ersten drei Punkte findet man eine allgemeine verfügbare und ausgetestete Technologie vor, wohingegen die letzten beiden Punkte die beiden größten Herausforderungen der aktuellen Forschung und Entwicklung darstellen. Man ist heute weit davon entfernt, über allgemein zugängliche und ausreichend genaue elektronische Karten zu verfügen. Ein damit verbundenes Problem, welches besonders Europa und vor allem Deutschland große Schwierigkeiten bereitet, ist der Umstand, daß die Daten in diesen Regionen von privatwirtschaftlichen Firmen erhoben und digitalisiert werden. Dadurch werden in Europa die Daten von mehreren

Firmen, wie EGT und TeleAtlas, parallel erhoben. Dies führt zur Etablierung verschiedener Datenstandards, was die Verwendbarkeit der Daten einschränkt. In den USA dagegen werden die Grunddaten vom "U.S. Bureau of Censor" zur Verfügung gestellt.

In besonderem Maße ist in den letzten Jahren die Entwicklung von Fahrernavigationssystemen vorangetrieben worden ([Dav89], [Mar89], [CCC<sup>+</sup>94], [BD95]). In großen europäischen (PROMETHEUS, DRIVE<sup>1</sup>), amerikanischen (IVHS<sup>2</sup>) und japanischen (VICS<sup>3</sup>) Projekten wird untersucht, wie Fahrer in einem Automobil im einzelnen und in der Gesamtheit so unterstützt werden können, daß sie entsprechend ihren Bedürfnissen in optimaler Weise zum Ziel gelangen. Diese Projekte bilden erst den Anfang allgemein einsatzfähiger Systeme. In sie wird sowohl von Seiten des Staates, als auch der Industrie viel Geld investiert. Beispielsweise wurde im Rahmen des *Pathfinder* Projektes (vgl. [Bad93]), an dem das kalifornische Verkehrsministerium (CALTRANS), das nationale Verkehrsministerium der USA (FHWA) und General Motors beteiligt sind, 25 von GM gespendete Wagen mit dem *TravelPilot* von BLAUPUNKT/BOSCH ausgerüstet. Der *TravelPilot* ist der Vorgänger des derzeit von BLAUPUNKT/BOSCH unter dem selben Namen vertriebenen Systems. Derzeit erwartet man den Einbau von einer Million Navigationssystemen in deutsche Fahrzeuge bis zum Jahr 2000 und einer Verdoppelung auf zwei Millionen bis zum Jahr 2002.

Das allgemeine Navigationsproblem läßt sich in acht Teilmodule unterteilen ([FHW93], [NCG93]):

1. Orientierung (Wo befindet man sich?)
2. Pfadsuche (Welcher Weg führt dort hin?)
3. Navigationshilfe (Wie kommt man dort hin?)
4. Weginformation (Was befindet sich entlang des Pfades?)
5. Pfadzustand (Wie ist der Weg?)
6. Fahrtenplanung (Wann kommt man an?)
7. Fahrtenzuteilung (Wer kommt dort als erstes an?)
8. Modusauswahl (Welche Fortbewegungsart wird gewählt?)

Jede dieser Teilmodule wird von verschiedenen Navigationssystemen in unterschiedlicher Weise verwendet. So wurden im Rahmen des IVHS-Programms 27 verschiedene Navigationsaufgaben im Kontext von Fahrernavigationssystemen klassifiziert (vgl. [CCC<sup>+</sup>94]):

1. *Information vor Beginn einer Fahrt*: Lieferung von Reiseinformation für einen kompletten Bereich von multimodalen Transport- und Beförderungsoptionen auf Nachfrage vor Beginn der Fahrt
2. *Fahrtbegleitende Fahrer-Information*: Aktuelle Verkehrsinformation für den Fahrer
3. *Fahrtbegleitende Reiseinformation*: Aktuelle Verkehrsinformation über multimodale Beförderungsoptionen während der Reise
4. *Reiseservice Information*: Information über den gesamten Bereich reiserelevanter Serviceleistungen und Einrichtungen auf Anfrage vor und während der Reise

---

<sup>1</sup>Dedicated Road Infrastructure for Vehicle Safety in Europe

<sup>2</sup>Intelligent Vehicle Highway Systems

<sup>3</sup>Vehicle Information and Communication System

5. *Fahrtenleitung*: Wegbeschreibungen im statischen und dynamischen Modus entsprechend benutzerspezifischer Einstellungen.
6. *Fahrtenidentifikation und -reservierung*: Information über Mitfahrangebote
7. *Unfall- und Behinderungs-Management*: Information über Verkehrsfluß und Umgebungszustände sowie Koordination bei Unfällen
8. *Reiseanforderungsmanagement*: Enthält Fahrbahnanforderungskontrolle, Parkplatz-Management-Systeme, Luftverschmutzung/Emissionsaufzeichnung, etc.
9. *Verkehrskontrolle*: Verkehrsüberwachung und -vorhersagen zur optimalen Steuerung des Verkehrsdurchflusses
10. *Elektronische Abrechnungssysteme*: Automatische Abbuchung von Gebühren von einem Konto für die Benutzung einer Straße
11. *Kommerzielle Fahrzeuginspektion*: Automatische Inspektion von Fahrzeug- und Fahrerinformation sowie Frachtgutscheinen
12. *Automatische Fahrzeuginspektionen am Wegesrand*: (vgl. mit verkürzten TÜV-Kontrollen)
13. *Kommerzielle Fahrzeugverwaltungsprozesse*: Elektronischer Erwerb von Transportgenehmigungen, automatische Aufzeichnung und finanztechnische Weitergabe von Kilometerständen und Verbrauch etc.
14. *Fahrzeuggestützte Sicherheitsüberwachung*: Überwachung des Zustands des Fahrzeugs, der Fracht und des Fahrers während der Fahrt.
15. *Kommerzielles Flotten-Management*: Identifikation, Lokalisation von Fahrzeugen, Fahrtenleitung, Flottenverwaltung und Echtzeitstatus zur Unterstützung der Flottenoperationen
16. *Öffentliches Verkehrsmittelmanagement*: Verwaltung von intermodalen, öffentlichen Verkehrsmitteln in Echtzeit, Serviceplanung, etc.
17. *Personalisierter öffentlicher Verkehr*: Auswertung und Nutzung von Kundenanfragen, Koordination von Fahrtenanfragen etc.
18. *Notfallbenachrichtigung und persönliche Sicherheit*: Benachrichtigung von verantwortlichen Stellen, Management von Einsatzplänen etc.
19. *Sicherheit öffentlicher Verkehrsmittel*: Überwachung öffentlicher Verkehrsmittel und Einrichtungen durch visuelle und auditive Sensoren
20. *Management von Einsatzfahrzeugen*
21. *Longitudinale Kollisionsvermeidung*: Auffahrunfallwarnung- und kontrolle, autonome bzw. kooperative und intelligente Fahrzeugsteuerung, Überholwarnung, etc.
22. *Laterale Kollisionsvermeidung*: Fahrbahnwechsel/Blinder Punkt Kollisionsvermeidung, Fahrbahnverfolgungskontrolle
23. *Kreuzungskollisionswarnung und -kontrolle*
24. *Kollisionsvermeidung durch Verbesserung der visuellen Sicht*: Aktive Verbesserung der visuellen Sicht zur Vermeidung von Kollisionen bei schlechten Sichtverhältnissen

25. *Übertretungsalarm*: Aktive Fahrzeug- und Straßenzustandsüberwachung zur Alarmierung des Fahrers
26. *Anwendung von Sicherheitsmaßnahmen unmittelbar vor einer Kollision*
27. *Vollautomatisierte Fahrzeugsteuerung*: Bietet ein vollauserüstetes automatisches Autobahnleitsystem und Kontrollfunktionen für die Infrastruktur und das Fahrzeug. Enthält die vorgestellten Kollisionsvermeidungssysteme, Übertretungsalarm und Funktionen zur Anwendung von Sicherheitsmaßnahmen unmittelbar vor einer Kollision.

Diese Anwendungen sind nicht strikt voneinander getrennt, sondern bauen teilweise aufeinander auf. Einige hören sich heutzutage noch sehr futuristisch an und überschreiten mit Sicherheit deutsche Datenschutzgesetze. Sie zeigen jedoch das allgemeine Potential von Navigationssystemen an und geben einen Hinweis darauf, was in Zukunft zumindest auf dem amerikanischen Markt angeboten werden wird. Beispielsweise sind testweise bereits zwei elektronische Abrechnungssysteme in den USA im Einsatz.

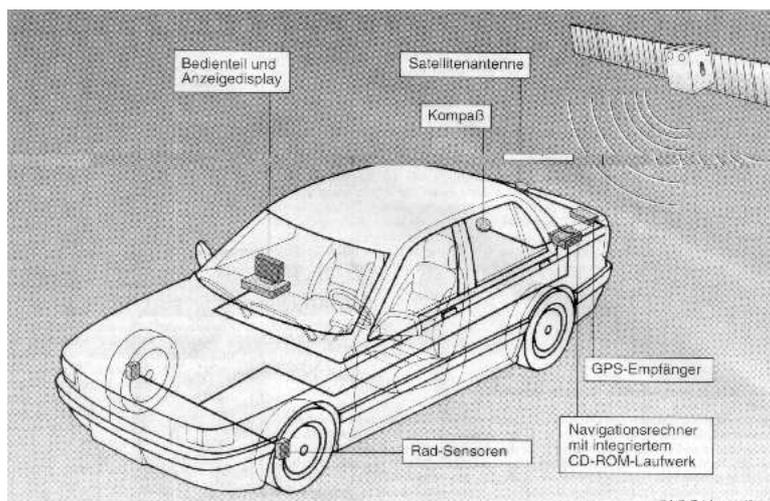
Jede einzelne Anwendung basiert auf den vorher indizierten Anforderungen in unterschiedlicher Weise (vgl. [NCG93], [CCC<sup>+</sup>94]). Vor diesem Hintergrund der Komplexität der Aufgaben, die im Kontext intelligenter Verkehrsleitsysteme auftreten, läßt sich ein inkrementelles Wegbeschreibungssystem, so wie es durch MOSES vorgestellt worden ist, als System zur Fahrtenleitung verstehen.

Eine natürliche technische Verwendung inkrementeller Wegbeschreibungen erfolgt im Zusammenhang mit intelligenten Fahrernavigationssystemen ([MWH93]). Von verschiedenen Firmen werden erste Fahrernavigationssysteme angeboten, die den Namen wirklich verdienen. So bieten BLAUPUNKT/BOSCH (TravelPilot) und PHILIPS (Carin) zwei sehr ähnliche Systeme an, die dem Fahrer über einen kleinen Monitor auf der Konsole seines Wagens die lokale Konfiguration der Straßenabschnitte anzeigt und Richtungsangaben mittels Pfeilsymbolen indiziert, sowie verbal eine kurze Beschreibung der nächsten Aktion am nächsten Verzweigungspunkt angibt (vgl. Abbildung ??). Im System von BLAUPUNKT/BOSCH wird die Position des Wagens durch Abgleich von raumbezogenen Daten vorgenommen, die via GPS empfangen, von einer CD abgerufen, sowie durch Radsensoren und per Kompaß ermittelt werden. Dies ist notwendig, da die Daten des GPS vom amerikanischen Verteidigungsministerium aus militärischen Gründen gestört werden, was zu einer Herabsetzung der Genauigkeit der Koordinaten führt. Dadurch erlauben die empfangenen Daten nur noch die Bestimmung der Lokation bis auf ca. 10 Meter genau.

Fahrernavigationssysteme Fahrernavigationssystem mit verschiedenen Informationskanälen (aus FAZ, 28. Mai 1996, Nr. 122, Seite T1) Navigation

Zur Präsentation von Navigationsdaten wird in prototypischen Systemen die Verwendbarkeit von sogenannten "Head-Up Displays" (HUD) überprüft, bei denen die Information auf die Windschutzscheibe direkt projiziert wird. Somit wird die Information direkt im natürlichen Sichtfeld des Fahrers angeboten ([BR92]). Über HUDs werden vor allem Geschwindigkeiten und Richtungen angezeigt. In zukünftigen Systemen soll dies um *künstliche Sichten* erweitert werden, um beispielsweise bei Nebel unter Einsatz von Infrarotsensoren ein genaueres Bild der Umgebung zu erlangen. Ebenso kann diese Technik dazu verwandt werden, um allgemeine Information über beispielsweise Gebäude, Lokationen und Veranstaltungen zu vermitteln.

Generell ist bei der Entwicklung von Navigationsschnittstellen darauf zu achten, daß der Fahrer nur relevante Information zur richtigen Zeit erhält ([FRH92]). Bedingt durch identische Prämissen kann MOSES in eingeschränkter Weise als Fahrernavigationssystem verstanden werden. MOSES erweitert momentan auf dem Markt befindliche Systeme dadurch, daß es über eine dynamische, inkrementelle Generierung von Beschreibungen verfügt und diese in Abhängigkeit



von der Situation präsentiert. Im Gegensatz zu Navigationssystemen, die sich momentan im Einsatz befinden, ist MOSES in der Lage Landmarken aus einer Situation in die Beschreibung zu integrieren. Zwar wird aktuell das Kartenmaterial um Landmarkeninformation erweitert, jedoch löst dies nicht das Problem, ob eine Landmarke überhaupt in der jeweiligen Situation sichtbar ist. Um Landmarken in solchen Navigationssystemen verwenden zu können ist es entweder notwendig, die Information über die Sichtbarkeit von Landmarken mit in der Karte zu speichern oder aber, wie in MOSES, dies über die Auswertung der visuellen Information zu lösen. Die Effizienz, die durch die erste Variante erreicht wird, geht zulasten der Flexibilität und Gültigkeit der Beschreibung. Möglicherweise kann eine Kombination aus beiden Ansätzen einen geeigneten Lösungsansatz bieten.

Einsatzfähige Systeme, wie der *TravelPilot* und *CARIN* verfügen über keinen direkten Zugang zur aktuellen Fahrsituation. Die einzige Information, die im erweiterten Sinne als situationsgebunden bezeichnet werden kann, ist die Bestimmung der Lokation des Fahrzeuges. Dadurch fehlt es den Systemen grundsätzlich an reaktiven Fähigkeiten, als auch der Möglichkeit, neue Information über Situationen zu erwerben. Es gibt zahlreiche Bestrebungen, Fahrernavigationssystemen mittels rundfunkbasierten, funkbasierten (z.B. SOCRATES (Europa), RACS (Japan) und ADVANCE (USA) in [RHA92]) oder barkenbasierten Technologien Verkehrsinformation zuzuführen. Aber selbst diese aufwendige Technologie erlaubt es dem Navigationssystem nicht immer, über eine ausreichende Repräsentation der aktuellen Situation zu verfügen. Zwischen der Übertragung einer Information zum Sender und von diesem zum Navigationssystem entsteht eine Informationslücke. Eine solche Lücke kann nur dadurch geschlossen werden, indem das Navigationssystem in ausreichender Weise individuell die aktuelle Situation analysiert. Im Idealfall muß ein Navigationssystem erkennen, wenn sich gerade ein Unfall entwickelt, um dem Fahrer entsprechende Handlungsvorschläge mitzuteilen. Weiterhin kann durch Rückkopplung des Navigationssystems zu zentralen Einrichtungen diese Information in einem globalen Rahmen analysiert und an andere Navigationssysteme weitergeleitet werden. Hierdurch wird ein Navigationssystem nicht nur als Empfänger, sondern auch als Sender verwendet, was die Kosten für extern angebrachte Sensoren verringert. Eine Analyse einer Situation ist jedoch äußerst komplex. Vor allem die Untersuchung visuell-optischer Daten ist vielfach untersucht worden. In experimentellen Umgebungen sind Modelle vorgeschlagen und implementiert worden, die ein Automobil autonom mit relativ hoher Geschwindigkeit auf einer markierten (z.B. [Enk93]) oder unmarkierten Fahrbahn (z.B. [THKS91]) navigieren. Solche Systeme setzen zahlreiche vereinfachende Annahmen voraus, so daß deren Einsatz im alltäglichen Straßenverkehr in kurz- und mittelfristiger Sicht ausgeschlossen erscheint.

Neben dem Einsatz im automobilen Verkehrsbereich wird untersucht, inwieweit sich Navigationssysteme in anderen Bereichen verwenden lassen können. Ein Navigationssystem bedingt grundsätzlich nur, daß sich der Rezipient real oder mental bewegt. Einen besonders großen Nutzen versprechen Navigationssysteme im Hinblick auf die Unterstützung von sehgeschädigten, insbesondere blinden Personen. Angewiesen auf haptische, olfaktorische und auditive Sinne sind Objekte, die sich durch diese Sinne nicht wahrnehmen lassen, für Blinde nicht zugänglich. Durch Navigationshilfen können solche Defizite für den Handlungsbedarf nahezu kompensiert werden. In experimentellen Systemen ist beispielsweise an der UC Santa Barbara mit relativ geringem technischen Aufwand (GPS, CD-ROM) ein tragbares Gerät entwickelt worden, welches verbale Beschreibungen von Objekten in der unmittelbaren Umgebung liefert ([LGK95], [GLK<sup>+</sup>91]). Dadurch hat der Benutzer nicht nur die Möglichkeit, seine unmittelbare Umgebung besser zu nutzen, sondern auch sein raumbezogenes Wissen langfristig zu erweitern, was seine Handlungsfähigkeiten erheblich verbessert. In ähnlicher Weise lassen sich derartige Navigationssysteme als Orientierungshilfe für komplexe Umgebungen konstruieren, wie z.B. für Flughäfen, Bahnhöfe, Museen und Krankenhäuser. Ein weites Gebiet eröffnet sich im Hinblick auf Rehabilitationsmaßnahmen von gehirnneurologisch geschädigten Personen, die nur noch über eingeschränkte räumliche Orientierungsfähigkeiten in ihren alltäglichen Umgebungen besitzen.<sup>4</sup> Der Einsatz von Navigationssystemen ermöglicht es Patienten mit bestimmten Schädigungen, ihre Umgebung teilweise wieder zu erlernen und die Abhängigkeit von Betreuern zu verringern, was nicht zuletzt ein finanzielles und personelles Problem ist ([HM94]).

In realen Umgebungen ist das Potential, welches durch Navigationssysteme im erweiterten Sinne gegeben ist, äußerst vielfältig. In erster Linie kommen Navigationssysteme im Verkehrsbereich zum Einsatz. Dabei entwickelte Technologien werden in Zukunft aber in vielfacher Weise verwendet werden, so auch als Orientierungs- und Zielfindungshilfe für Fußgänger im Alltag und beispielsweise in Museen und Supermärkten. Im Hinblick auf den Spruch „Zeit ist Geld“ besitzen Navigationssysteme gute Aussichten, dort eingesetzt zu werden, wo sie in ausreichendem Maße dazu beitragen, funktionale Probleme zu lösen. Dazu ist es notwendig, daß Navigationssysteme auf immer mehr Daten über den Problemraum zugreifen können. Da eine vollständige Repräsentation einer realen Umgebung weder möglich noch wünschenswert ist, wird es nach wie vor nur zu suboptimalen Lösungen kommen. Navigationssysteme werden sich für räumliche Probleme zu dem entwickeln, was Taschenrechner heutzutage für numerische Probleme sind. Es ist nur zu hoffen, daß die Fähigkeit des Menschen, raumbezogene Probleme zu lösen, nicht in gleicher Weise abnimmt, wie dies beim Kopfrechnen der Fall ist. Besonders im Hinblick auf Navigationssysteme besteht eines der Probleme darin, geeignete Kommunikationsweisen zwischen Mensch und Maschine zu finden. Neben unimodalen Systemen gibt es bereits Modelle, die untersuchen, inwieweit multimodale Ein- und Ausgaben die Effektivität der Kommunikation steigern können. Beispielsweise wird im System *MOFA* in einem Wegbeschreibungskontext untersucht, in welcher Weise Gesten und Natürliche Sprache miteinander gekoppelt werden können (vgl. [SK96]).

Neben dem Einsatz von Navigationssystemen bezüglich realen Umgebungen besitzen sie auch Anwendungspotential in virtuellen Umgebungen. Dieses soll im weiteren kurz diskutiert werden.

### 5.2.2 Navigationshilfe in virtuellen Umgebungen

In virtuellen Umgebungen werden raum-zeitliche Metaphern verwendet, um dem Benutzer einen möglichst natürlichen Zugang zu den Daten bzw. zur Funktionalität des Systems zu erlauben. Beispielsweise ist der *Mülleimer* der MacOS-Oberfläche mehr als nur ein Symbol. Um eine Datei zu löschen, „nimmt“ der Benutzer sie auf, „geht“ zum Mülleimer und „wirft“ die Datei

---

<sup>4</sup>Für einen Überblick entsprechender Forschungsarbeiten zum Thema *Neglekt* sei auf die Sonderausgabe der Zeitschrift *Neurophysiological Rehabilitation* verwiesen ([HM94]).

hinein. Ganz natürlich werden dem Mülleimer-Symbol automatisch eine inhärente Bedeutung sowie Prozeduren zugeordnet. In gleicher Weise werden in virtuellen Umgebungen Entitäten via Metaphern Bedeutungen und Verwendungsweisen zugeordnet. Die Verwendung von Entitäten in virtuellen Welten hängt entscheidend von der Natürlichkeit ab, mit der Metaphern des alltäglichen Lebens mit Entitäten verknüpft werden. Beispielsweise wird auf dem "world-wide-web" (WWW) versucht, durch Verwendung von „Haus“- und „Stadt“-Metaphern einzelne WWW-Seiten zu integrieren, bzw. zu strukturieren. Der Übergang von einer Seite zu einer anderen wird dadurch mittels einer „Pfad“-Metapher verständlich. Verbindungen zwischen Seiten unterliegen generell keinen Beschränkungen. Jede Seite kann prinzipiell mit jeder anderen verknüpft werden. Wie sich aber herausstellt, verwirrt dies den Benutzer mehr, als es ihn unterstützt. Was dem Benutzer fehlt, ist eine geeignete Orientierung. Aus diesen Gründen werden Metaphern dazu verwendet, den Umgang mit virtuellen Umgebungen zu verbessern. Jedoch werden nicht alle realen Handlungsweisen übernommen. Beschränkungen beispielsweise, die durch physikalische Gesetze gegeben werden, werden im WWW nicht integriert bzw. bewußt aufgehoben. Einer der größten Unterschiede ist, daß die Distanz zwischen zwei WWW-Seiten nur durch die Geschwindigkeit des Netzes und der beteiligten Rechner bestimmt ist. Dies führt zu scheinbar paradoxen Situationen, in denen der Übergang von der Heim-Seite des Benutzers zu einer Seite in San Francisco schneller erfolgt, als der zu einer Seite einer Person im direkten regionalen Umkreis. Personen, die real nahe erscheinen, sind virtuell weiter entfernt und vice versa. Einerseits ist dies ein Vorteil, andererseits müssen sich Benutzer an diese Metriken gewöhnen. Da die Verwendung virtueller Welten durch eine größere Anzahl von Personen weltweit gerade erst begonnen hat, ist es heute noch nicht abzusehen, inwieweit Metaphern des alltäglichen Lebens nur als Übergang oder als wesentlicher Bestandteil im Umgang mit virtuellen Umgebungen verwendet werden. Da der Erfolg von virtuellen Umgebungen wesentlich von der Einfachheit abhängt, mit der man diese benutzen kann, ist zu vermuten, daß Metaphern, die bezüglich realer Umgebungen gelten, verstärkt eingesetzt werden. Dabei wird es zu Erweiterungen der Bedeutung einzelner Metaphern kommen, so daß sie mit der Zeit immer mehr von ihrer ursprünglichen Bedeutung verlieren werden. Der MacOS-Mülleimer ist ein Beispiel hierfür. „Wirft“ man eine Diskette in den Mülleimer, so wird diese aus dem Diskettenschacht ausgeworfen. In realen Umgebungen erwartet man ein solches Verhalten von einem Mülleimer nicht.

Wegbeschreibungen besitzen zahlreiche Metaphern, die in virtuellen Umgebungen verwendet werden. So lassen sich in virtuellen Umgebungen Landmarken, Verzweigungspunkte, Pfade, Start- und Ziellokationen wiederfinden. Die Metapher einer Landmarke wird größtenteils besonders interessanten und salienten Lokationen<sup>5</sup> zugeordnet. Da ein Benutzer nur eine Lokation zu einem Zeitpunkt lesen kann, bietet eine Lokation verschiedene Verzweigungsmöglichkeiten an, über die man zu neuen Lokationen gelangen kann. Somit kann eine Lokation auch mittels einer Verzweigungspunkt-Metapher verstanden werden. Verbindungen zwischen Lokationen werden mittels Pfad-Metaphern erfaßt. Interessanterweise wird das „Umherwandern“ im WWW als „surfen“ bezeichnet, was daher rührt, daß der Benutzer meist ziellos von einer Seite zur nächsten übergeht. Dennoch ist der Bewegungscharakter, der durch surfen indiziert wird, offensichtlich. Start- und Ziellokationen sind i.a. gleichbedeutend mit einzelnen Lokationen der virtuellen Umgebung, was z.B. im WWW eine WWW-Seite ausmacht. Somit lassen sich alle grundlegenden Elemente in virtuellen Umgebungen identifizieren, die dem WWW ähnlich sind. Zu vermuten ist, daß ein Benutzer sich in Zukunft aus dem WWW eigene Umgebungen in komplexer Weise zusammensetzen kann.<sup>6</sup>

Im Hinblick auf die kommerzielle Nutzung solcher Technologien zeigt sich, daß herkömmliche

---

<sup>5</sup>Bezogen auf das WWW sind Lokationen am ehesten mit WWW-Seiten zu unifizieren.

<sup>6</sup>Die Verwaltung von WWW-Seiten in strukturierten Listen, wie in Netscapes Navigator 2.0, ist hierzu nur ein Anfang.

Marketing-Strategien auf virtuelle Umgebungen mehr oder weniger direkt übertragen werden. In virtuellen Kaufhäusern beispielsweise wird der Kunde in Zukunft in der gleichen Weise wie in realen Kaufhäusern an Produktangeboten vorbeigeführt, bevor er das Produkt erreicht, welches er eigentlich beabsichtigt zu kaufen. Zusätzlich wird der Kunde gezielter auf Eigenschaften von bestimmten Produkten aufmerksam gemacht. Die „Pfade“, die der Benutzer während des Einkaufs verwendet hat, werden gespeichert und ausgewertet, damit auf dieser Datenbasis effizientere Marketing-Strategien entwickelt werden können. Betritt der Konsument ein virtuelles Kaufhaus, so wird er möglicherweise einen elektronischen Einkaufszettel abgeben. Dieser wird verwendet, um einen optimalen Einkaufspfad zu bestimmen. Die Optimalität eines Pfades wird dabei sowohl dadurch bestimmt, daß der Benutzer nicht den Eindruck gewinnt, sich unnötig lange im Kaufhaus zu befinden, und daß er nach marketing-strategischen Gesichtspunkten an Zusatzprodukten und Marktneuheiten entlanggeführt wird. Aktuell ist der Übergang von einer WWW-Seite zur nächsten statisch. In Zukunft werden solche Übergänge dynamisch und „on the fly“ animiert werden, so daß beim Benutzer der Eindruck einer Bewegung erzeugt wird. Es ist aber abzusehen, daß der Benutzer in gleicher Weise durch Informationsüberflutung überlastet werden wird, wie es heutzutage bereits in realen Umgebungen und heutigen Internet-Anwendungen der Fall ist. Eine der schwierigsten Aufgaben wird es sein, dem Benutzer nur soviel Information über Dinge zu liefern, die er einerseits wirklich verlangt und andererseits letztendlich als Zusatzinformation erachtet.

Vor diesem Hintergrund sind Wegbeschreibungen im erweiterten Sinne von hervorragendem Interesse. Alle Elemente, aus denen sich Wegbeschreibungen konstituieren, finden sich i.a. in ausreichend komplexen, virtuellen Welten wieder. Die metaphorische Verwendung dieser Elemente ist zwar in gewisser Weise an die jeweilige Umgebung anzupassen, verliert jedoch nicht die grundlegende Bedeutung. Durch Anpassung der Elemente läßt sich MOSES als Navigationssystem in einem virtuellen Kaufhaus oder dem WWW allgemein verwenden. Information über Landmarken kann dazu verwendet werden, dem Benutzer die aktuelle virtuelle Lokation zu beschreiben. Verzweigungspunkte können einerseits im obig beschriebenen Sinne eine Reise durch ein Kaufhaus beschreiben. Andererseits kann dem Benutzer auch kurz beschrieben werden, was sich hinter den Verzweigungen befindet.

### 5.3 Erweiterungsmöglichkeiten

Abschließend soll aufgezeigt werden, in welcher Weise das vorgeschlagene Modell unmittelbar, aber auch mittelbar erweitert werden kann.

#### 1. Hierarchische Beschreibungen:

In dem vorgeschlagenen Modell werden inkrementelle Wegbeschreibungen behandelt, die die unmittelbar folgende Aktion am nächsten Verzweigungspunkt indizieren. Wie der Korpus zeigt, gaben die Probanden in seltenen Fällen hierarchisierte Beschreibungen: „Jetzt gehen wir zum Bahnhof. Dazu gehen wir erstmal geradeaus.“ Zu Beginn einer solchen Beschreibung wird typischerweise ein lokales Ziel, welches meist durch eine Landmarke gegeben ist, indiziert, bevor der lokale Pfad dorthin beschrieben wird. Um solche Beschreibungen geben zu können, benötigt der Sprecher mehr räumliches Wissen über die Umgebung, als über Perzeptionsprozesse extrahiert werden kann. Aus diesem Grunde ist dieses Phänomen in der vorliegenden Arbeit nicht berücksichtigt worden. Es stellt jedoch eine natürliche Erweiterung dar.

#### 2. Pfadrelationen

Wie in Kapitel 2.3.2 einleitend erwähnt, wird zwischen topologischen, projektiven und

pfadbezogenen, räumlichen Relationen unterschieden. In dem Modell wurden die ersten beiden Relationenarten verwendet, um eine räumliche Vorstellung über eine Umgebung aufzubauen. Pfadrelationen erweitern diese Vorstellung dahingehend, daß ein Objekt durch eine Trajektorie abstrahiert wird, der eine räumliche Bedeutung zugeordnet wird. Bei Abbiegevorgängen wird an Verzweigungspunkten den Objekten keine Trajektorie, sondern eine statische Punkt-zu-Punkt-Beziehung zugeordnet, wie sie durch topologische und projektive räumliche Relationen repräsentiert werden. Pfadrelationen werden auf sprachlicher Ebene vor allem für unterstützende Beschreibungen während des Zeitrahmens  $t_1$  sprachlich indiziert, wie zum Beispiel in der folgenden Beschreibung: „Gehen Sie an der Kirche entlang.“

### 3. Bewegungsverben

Mit einer Modellierung von Pfadrelationen eng verbunden ist eine solche für Bewegungsverben (vgl. [Kau89]).<sup>7</sup> In dem dieser Arbeit zugrunde liegenden Korpus sind ausschließlich die beiden Verben „gehen“ und „abbiegen“ verwendet worden. „Abbiegen“ wird im Kontext der inkrementellen Wegbeschreibungen dann verwendet, wenn es sich um einen Abbiegevorgang handelt. Entsprechend wird „gehen“ nahezu ausschließlich für die Beschreibung einer translatorischen Bewegung verwendet. Im Hinblick auf allgemeine Raumbeschreibungen ist eine solche approximative Repräsentation zu verfeinern.

### 4. Kognitive Karten

In dieser Arbeit ist ausgeschlossen worden, daß der Sprecher über ein räumliches Langzeitgedächtnis der Umgebung, in der er sich bewegt, verfügt. Das räumliche Wissen, welches er zur Verfolgung des Weges benötigt, extrahiert er aus topographischen Netzrepräsentationen. Information über Landmarken gewinnt er durch Analyse der Situation. In vielen Fällen besitzt der Sprecher eine mehr oder weniger gut ausgebildete Kognitive Karten der Umgebung. *Räumliche Konfigurationsbeschreibungen* repräsentieren einen Teil des räumlichen Wissens, welches für den Aufbau von Kognitiven Karten notwendig ist. Aus dieser Perspektive heraus, kann eine räumliche Konfigurationsbeschreibung als Eingabe für Prozesse zur Generierung Kognitiver Karte verstanden werden. In weiteren Arbeiten muß untersucht werden, welches weitere räumliche Wissen berücksichtigt und miteinander integriert werden muß, um eine räumliche Repräsentation, ähnlich einer Kognitiven Karte, aufbauen zu können.

### 5. Sprachgesteuerte Aufmerksamkeitsfokussierung

Der Rezipient einer inkrementellen Wegbeschreibung richtet seine Aufmerksamkeit nicht notwendigerweise in der gleichen Weise wie der Sprecher aus. In solchen Fällen versucht der Agent die Aufmerksamkeit des Rezipienten durch sprachliche oder gestische Hinweise zu lenken. Um eine entsprechende sprachliche Beschreibung generieren zu können, muß der Agent ein Hörermodell aufbauen. Mittels diesem vergleicht der Agent die Ausrichtung des angenommenen egozentrischen Referenzsystems des Rezipienten mit der intendierten. An Hand der räumlichen Abweichung lassen sich geeignete Raumbeschreibungen generieren.

### 6. Vollständiger Sprecher-Hörer-Dialog und Dialoggedächtnis

Bedingt durch die Fokussierung von Beschreibungen an Verzweigungspunkten, ist der Aspekt der Kommunikation in dieser Arbeit in den Hintergrund getreten. Wegbeschreibungen sind allgemein jedoch dialogisch. Damit ist gemeint, daß der Rezipient nachfragt, um eine dediziertere Beschreibung zu erhalten oder aber um sich zu versichern, daß er noch auf dem richtigen Weg ist. In solchen dialogischen Kontexten muß der Sprecher über

---

<sup>7</sup>Einen möglichen Ansatz zur Modellierung von Bewegungsverben findet sich in [Mai90].

ein Dialoggedächtnis verfügen, um einerseits die Aussagen des Rezipienten zu verstehen und um andererseits auf Vorhergesagtes referenzieren zu können. Weiterhin ist ein geeignetes Benutzermodell zu fordern, welches es dem Sprecher erlaubt, in geeigneter Weise auf die (Nach-) Fragen des Rezipienten zu reagieren. Dies bedingt eine Hierarchisierung des räumlichen Wissens und geeigneter Zugriffsprozessen. Beispielsweise ist der Weg für einen Einheimischen anders zu beschreiben, als für einen Auswärtigen.

## 7. Emotionen

Ein Aspekt, der in den meisten KI Modellen nicht behandelt wird, ist der der Emotionen. Wie empirische Studien im Kontext von Fahrernavigationssystemen zeigen, sind Emotionen wichtige Indikatoren für die Inhaltsbestimmung und die Präsentation einer Beschreibung.<sup>8</sup> Ein Indikator ist beispielsweise die Art und Weise, mit der das Gaspedal des Autos bedient wird. In Stressituationen erfolgt die Bedienung stoßartiger als in entspannten Situationen. Davon läßt sich ableiten, daß der Rezipient, in diesem Fall der Fahrer, nicht durch unnötige Information belastet werden möchte. In Bezug auf das vorgestellte Modell erfolgt dies durch eine Reduzierung der Beschreibungen bezüglich der Zeitrahmen  $T_4$  und  $T_5$ . In welcher Weise solche Faktoren insbesondere Raumbeschreibungen beeinflussen, muß in weiteren empirischen Untersuchungen analysiert werden.

## 8. Anwendung auf andere Domänen

Eine der Prämissen dieser Arbeit ist, daß Konzepte, mit denen Raumbeschreibungen analysiert werden, in verschiedenen Domänen überprüft werden müssen. In der vorliegenden Arbeit sind es stadtähnliche Umgebungen gewesen, bezüglich derer bestehende Konzepte angepaßt und verfeinert worden sind. Vor allem die Erweiterungen bezüglich Referenzsystemen müssen in anderen Domänen überprüft werden. Andere Domänen, die im Rahmen des Projektes REAL verwendet werden, sind die Domänen eines Bahnhofs und eines Supermarktes.

Abschließen möchte ich diese Arbeit mit einem Zitat von Steven Vere und Timothy Bickmore (vgl. [VB90]).

“At one time it was thought that intelligence and AI had something to do with generality. The agent is subject to criticism from researchers specializing in one intelligence component, who may be able to point out that their one component works better than the corresponding component of an integrated agent. However, agents deserve credit for their “horizontal depth”. Like an athlete in the decathlon, an agent should be able to earn points for its performance in each cognitive event, without having to beat all of the narrower systems in their one speciality event. ... Researches specializing in a particular component of intelligence should realize that an integrated agent is in fact a validation and justification of their own work.”

---

<sup>8</sup>Diesen interessanten Hinweis verdanke ich Xavier Briffault.

# Literaturverzeichnis

- [ABHR86] E. André, G. Bosch, G. Herzog, and T. Rist. Characterizing trajectories of moving objects using natural language path descriptions. In *Proc. of the 7<sup>th</sup> ECAI, Vol. 2*, volume 2, pages 1–8, Brighton, UK, 1986.
- [ABHR87] E. André, G. Bosch, G. Herzog, and T. Rist. Coping with the intrinsic and the deictic uses of spatial prepositions. In K. Jorrand and L. Sgurev, editors, *Artificial Intelligence II: Methodology, Systems, Applications*, pages 375–382. North-Holland, Amsterdam, 1987.
- [Abk82] G. G. Abkarian. Comprehension of deictic locatives: The object "behind" it. *Journal of Psycholinguistic Research*, 11:229–245, 1982.
- [Acr88] L. Acredolo. Infant mobility and spatial development. In Stiles-Davis, M. Kritchevsky, and U. Bellugi, editors, *Spatial cognition: Brain bases and development*, pages 157–166. Erlbaum, Hillsdale, NJ, 1988.
- [AHR89] E. André, G. Herzog, and T. Rist. Natural language access to visual data: Dealing with space and movement. In F. Nef and M. Borillo, editors, *Logical Semantics of Time, Space and Movement in Natural Language. Proc. of 1<sup>st</sup> Workshop*. Hermès, 1989.
- [AK85] G. L. Allen and K. C. Kirasic. Effects of the cognitive organization of route knowledge on judgments of macrospatial distance. *Memory and Cognition*, 13(3):218–227, 1985.
- [AK87] D. E. Appelt and A. Kronfeld. A computational model of referring. In *Proc. of the 10<sup>th</sup> IJCAI*, pages 640–647, Milan, Italy, 1987.
- [AK93] A. Abella and Kendler. Qualitatively describing objects using spatial prepositions. In *Proc. of the 11<sup>th</sup> AAAI*, pages 536–540, Washington, DC, 1993.
- [AK96] B. Andes and R. Knop. Ein wissensbasierter inkrementeller suchalgorithmus auf straßenkarten. Endbericht des Fortgeschrittenenpraktikums *Inkrementelle Wegbeschreibungen*, 1996.
- [All81a] G. L. Allen. A developmental perspective on the effects of "subdividing" macro spatial experience. *Journal of Experimental Psychology: Human Learning and Memory*, 7:120–132, 1981.
- [All81b] J. F. Allen. A general model of action and time. Technical Report 97, Computer Science Department, University of Rochester, NY, 1981.
- [All87] G. L. Allen. Cognitive influence on the acquisition of route knowledge in children and adults. In P. Ellen and C. Thinus-Blanc, editors, *Cognitive Processes and Spatial*

- Orientation in Animal and Man*, volume 2: Neurophysiological and Developmental Aspects of *NATO ASI Series, Series D: Behavioural and Social Sciences*, pages 274–283. Martinus Nijhoff, Dordrecht, 1987.
- [ALM64] D. Appleyard, K. Lynch, and J. Myer. *The View from the Road*. MIT Press, Cambridge, MA, 1964.
- [And83] J. R. Anderson. *The architecture of cognition*. Harvard University Press, Cambridge, MA, 1983.
- [And88] E. André. Generierung natürlichsprachlicher äusserungen zur simultanen beschreibung von zeitveränderlichen szenen: Das system soccer. Memo 26, Universität des Saarlandes, SFB 314 (VITRA), 1988.
- [Ans88] L. Anselin. *Spatial econometrics: Methods and models*. Martinus Nijhoff, Dordrecht, 1988.
- [APO75] L. P. Acredolo, H. L. Pick, and M. G. Olsen. Environmental differentiation and familiarity as determinants of children's memory for spatial location. *Developmental Psychology*, 11:495–501, 1975.
- [App69] D. Appleyard. Why buildings are known. *Environment and Behavior*, 1:131–156, 1969.
- [App70] D. Appleyard. Styles and methods of structuring a city. *Environment and Behavior*, 1:131–156, 1970.
- [App85a] D. E. Appelt. Planning english referring expressions. *Artificial Intelligence*, 26(1):1–33, 1985.
- [App85b] D. E. Appelt. *Planning English Sentences*. Cambridge University Press, Cambridge, London, 1985.
- [App85c] D. E. Appelt. Some pragmatic issues in the planning of definite and indefinite noun phrases. In *Proceedings of the 23th Annual Meeting of the Association for Computational Linguistics*, pages 198–203, 1985.
- [Arn74] R. Arnheim. *Art and visual perception*. University of California, Berkeley, 1974.
- [ASR78] G. Allen, A. Siegel, and R. Rosinski. The role of perceptual context in structuring spatial knowledge. *Journal of Experimental Psychology: Human Learning and Memory*, 4(6):617–630, 1978.
- [B34] K. Bühler. *Sprachtheorie*. Fischer, Jena, 1934.
- [Bad93] A. S. Badillo. Transportation and navigation. In G. H. Castle III, editor, *Profiting from a Geographic Information System*, pages 161–176. GIS World, Inc., Fort Collins, CO, 1993.
- [Bau95] J. Baus. Moses-speak: Objektauswahl basierend auf visueller salienz in dynamischen 3d-umgebungen. Master's thesis, Fachbereich Informatik, Universität des Saarlandes, 1995. in Vorbereitung.
- [BB93] X. Briffault and A. Braffort. Towards a model of cooperation between natural language and natural gestures to describe spatial knowledge. In *PACLING'93: Transcending language boundaries*, Vancouver, Canada, 1993.

- [BD95] X. Briffault and M. Denis. Analyse d'un corpus de dialogues de navigation à bord d'un véhicule automobile. Rapport Interne LIMSI 95-28, LIMSI, Orsay, 1995.
- [Bie67] M. Bierwisch. Some semantic universals of german adjectivals. *Foundations of Language*, 3:1–36, 1967.
- [Bie90] I. Biederman. Higher-level vision. In D. N. Osherson, S. M. Kosslyn, and J. M. Hollerbach, editors, *Visual Cognition and Action: An Invitation to Cognitive Science ( Volume 2)*, pages 41–72. MIT Press, Cambridge, MA, 1990.
- [Bin71] T. O. Binford. Visual perception by computer. In *Proc. IEEE Conf. on Systems and Control*, 1971.
- [BJKZ85] R. Bajcsy, A. Joshi, E. Krotkov, and A. Zwarico. Landscan: A natural language and computer vision system for analyzing aerial images. In *Proc. of the 9<sup>th</sup> IJCAI*, pages 919–921, Los Angeles, CA, 1985.
- [BL87] M. Bierwisch and E. Lang. Etwas länger - viel tiefer - immer weiter: Epilog zum dimensionsadjektivenprojekt. In M. Bierwisch and E. Lang, editors, *Grammatische und konzeptuelle Aspekte von Dimensionsadjektiven*, pages 649–699. Akademie-Verlag (Studia Grammatica 26/27), Berlin, 1987.
- [Bla91] M. Blades. Wayfinding theory and research: The need for a new approach. In D. M. Mark and A. U. Frank, editors, *Cognitive and Linguistic Aspects of Geographic Space*, pages 137–165. Kluwer, Dordrecht, 1991.
- [BMB70] J. M. Blaut, G. F. McCleary, and A. S. Blaut. Environmental mapping in young children. *Environment and Behavior*, 2:335–349, 1970.
- [BNH86] B. Bürkle, H. Nirmaier, and T. Herrmann. "von dir aus ..." zur hörerbegrenzten lokalen referenz. Bericht 10, Forschergruppe Sprechen und Sprache im sozialen Kontext, Heidelberg/Mannheim, 1986.
- [Bol63] O. F. Bollnow. *Mensch und Raum*. Kohlhammer, Stuttgart, 1963.
- [BP83] J. Barwise and J. Perry. *Situations and Attitudes*. MIT Press, Cambridge, MA, 1983.
- [BR92] H. Bubb and G. Reichart. *Braucht man HUD's im Kraftfahrzeug? Ergebnisse einer Literaturrecherche und eigener Versuche*, volume 948 of *VDI-Gesellschaft Fahrzeugtechnik*, pages 101–125. VDI Verlag, Düsseldorf, 1992.
- [Bra81] M. Bratman. Intention and means-end reasoning. *The Philosophical Review*, XC(2):252–265, 1981.
- [Bra84] M. Bratman. Two faces of intention. *The Philosophical Review*, XCIII(3):375–405, 1984.
- [Bre92] B. Brehmer. Dynamic decision making: Human control of complex systems. *Acta Psychologica*, 81:211–241, 1992.
- [Bro58] D. E. Broadbent. *Perception and communication*. Pergamon Press, 1958.
- [Bro81] R. A. Brooks. Symbolic reasoning among 3-d models and 2-d images. *Artificial Intelligence*, 17:285–348, 1981.

- [Bro86] R. A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal Robotics and Automation*, RA2:14–23, 1986.
- [Bry92] D. J. Bryant. A spatial representation system in humans. *Journal of Memory and Language*, 31:74–98, 1992.
- [BS83] R. Byrne and E. Salter. Distances and directions in the cognitive maps of the blind. *Canadian Journal of Psychology*, 37:293–299, 1983.
- [BS85] R. J. Brachman and J. G. Schmolze. An overview of the kl-one knowledge representation system. *Cognitive Science*, 9(2):171–216, 1985.
- [BS95] A. Blocher and J. R. J. Schirra. Optional deep case filling and focus control with mental images: Antlima-koref. In *Proc. of the 14<sup>th</sup> IJCAI*, pages 19–25, Montreal, Canada, 1995.
- [BTF92] D. J. Bryant, B. Tversky, and N. Franklin. Internal and external spatial frameworks for representing described scenes. *Journal of Memory and Language*, 31:74–98, 1992.
- [BW76] R. J. Beck and D. Wood. Cognitive transformation of information from urban geographic fields to mental maps. *Environment and Behavior*, 8:199–238, 1976.
- [Cad79] M. Cadwallader. Problems in cognitive distance: Implications for cognitive mapping. *Environment and Behavior*, 11:559–576, 1979.
- [Car91] H. Carstensen. The complexity of testing the equivalence of transition sequences. In G. Rozenberg, editor, *Advances in Petri Nets 1991*, pages 48–57. Springer, Berlin, Heidelberg, 1991.
- [Car93] M. Carroll. Deictic and intrinsic orientation in spatial descriptions: A comparison between english and german. In J. Altarriba, editor, *Cognition and culture: A cross-cultural approach to cognitive psychology*, page 1993. Elsevier, Amsterdam, 1993.
- [Cas57] E. Cassirer. *The philosophy of symbolic forms: The phenomenology of knowledge*, volume 3. Yale University Press, New Haven, 1957.
- [CCC<sup>+</sup>94] R. Church, D. Coughlan, Th. Cova, M. Goodchild, J. Gottsegen, and D. Lemberg. Final report: Caltrans agreement 65t155. Technical Report 94-6, National Center for Geographic Information and Analysis (NGIA), UC Santa Barbara, 1994.
- [CGGT87] H. Couclelis, R. G. Golledge, N. Gale, and W. R. Tobler. Exploring the anchor point hypothesis of spatial cognition. *Journal of Environmental Psychology*, 7:99–122, 1987.
- [Cha82] W. G. Chase. Spatial representations of taxi drivers. In D.R. Rogers and J.A. Sloboda, editors, *Acquisition of Symbolic Skills*. Plenum, New York, 1982.
- [Cha87] D. Chapman. Planning for conjunctive goals. *Artificial Intelligence*, 32(3):333–377, 1987.
- [CJ85] I. Carsten and T. Janson. Verfahren zur evaluierung räumlicher präpositionen anhand geometrischer szenenbeschreibungen. Master's thesis, Fachbereich für Informatik, Universität Hamburg, 1985.
- [CL85] R. Chatila and J.-P. Laumond. Position referencing and consistent world modeling for mobile robots. In *International Conference on Robotics and Automation*, pages 138–170. IEEE, 1985.

- [Cla73] H. H. Clark. Space, time, semantics, and the child. In T. E. Moore, editor, *Cognitive Development and the Acquisition of Language*. Academic Press, New York, 1973.
- [CMC94] D. N. Chin, M. McGranaghan, and T.-T. Chen. Understanding location descriptions in the lei system. In *Proc. of the 4th. Conference on Applied Natural Language Processing*, pages 138–143, Stuttgart, Germany, 1994.
- [CMD<sup>+</sup>91] M. Corbetta, F. M. Miezin, S. Dobmeyer, G. L. Shulman, and S. E. Peterson. Selective and divided attention during visual discrimination of shape, color, and speed: Functional anatomy by positron emission tomography. *The Journal of Neuroscience*, 11:2383–2402, 1991.
- [CMP<sup>+</sup>95] Th. Clermont, C. Meier, M. Pomplun, E. Prestin, H. Rieser, H. Ritter, and B. Velichkovsky. Augenbewegung, fokus und referenz. Technical Report 8, Situierete Künstliche Kommunikatoren, SFB360, Universität Bielefeld, 1995.
- [Coh84] P. R. Cohen. The pragmatics of referring and the modality of communication. *Computational Linguistics*, 10:97–146, 1984.
- [Coh87] A. G. Cohn. A more expressive formulation of many sorted logic. *Journal of Automated Reasoning*, 3:113–200, 1987.
- [Cor82] M. C. Corballis. Mental rotation: Anatomy of a paradigm. In M. Potegal, editor, *Spatial abilities*, page 1982. Academic Press, New York, 1982.
- [Cou95] H. Couclelis. Verbal directions for way-finding: space, cognition, and language. In J. Portugali, editor, *The Construction of Cognitive Maps*. Kluwer, Dordrecht, 1995. in print.
- [Cow79] A. Cowey. Cortical maps and visual perception. the grindley memorial lecture. *Quarterly Journal of Experimental Psychology*, 31:1–17, 1979.
- [Cox81] M. V. Cox. Interpretation of the spatial prepositions 'in front of' and 'behind'. *International Journal of Behavioral Development*, 4:359–368, 1981.
- [CRI93] L. A. Carlson-Radvansky and D. E. Irwin. Frames of reference in vision and language: Where is above? *Cognition*, 46:223–244, 1993.
- [CS69] S. Carr and D. Schissler. The city as a trip: Perceptual selection and memory in the view from the road. *Environment and Behavior*, 1:7–36, June 1969.
- [CWG86] H. H. Clark and D. Wilkes-Gibbs. Referring as a collaborative process. *Cognition*, 22:1–39, 1986.
- [Dav86] E. S. Davis. *Representing and Acquiring Geographic Knowledge*. Pitman, London, 1986.
- [Dav89] J. R. Davis. *Back Seat Driver: Voice Assisted Automobile Navigation*. PhD thesis, Media Arts and Science Section, MIT, Cambridge, MA, 1989.
- [DB87] T. Dean and M. Boddy. Incremental causal reasoning. In *Proc. of AAAI-87*, pages 196–201, Seattle, WA, 1987.
- [DB88] T. Dean and M. Boddy. An analysis of time-dependent planning. In *Proc. of AAAI-88*, pages 49–54, St. Paul, MN, 1988.

- [Del80] G. S. Dell. *Phonological and lexical encoding in speech production: An analysis of naturally occurring and experimentally elicited speech errors*. PhD thesis, University of Toronto, 1980.
- [Der91] G. Derefeldt. Colour appearance systems. In P. Gouras, editor, *The Perception of Colour*, volume 6, pages 218–261. Macmillan, London, 1991.
- [Dev76] A. S. Devlin. The small town cognitive map: Adjusting to a new environment. In G. Moore and R. G. Golledge, editors, *Environmental Knowing: Theories, Research and Methods*. Hutchinson and Ross, Stroudsburg, PA, 1976.
- [DIN80] DIN 6164. *System der DIN-Farbenkarte für den 2°-Normalbeobachter*, volume Teil 1: DIN-Farbenkarte. Beuth-Verlag, Berlin, 1980.
- [DL88] R. M. Downs and L. S. Liben. Through a map darkly: Understanding maps as representations. *The Genetic Epistemologist*, 16:11–18, 1988.
- [DP82] W. Deutsch and T. Peschmann. Social interaction and the development of definite descriptions. *Cognition*, 11:155–168, 1982.
- [DS73] R. M. Downs and D. Stea. Cognitive maps and spatial behaviour: Process and products. In R. M. Downs and D. Stea, editors, *Image and Environment. Cognitive Mapping and Spatial Behaviour*, pages 8–26. Aldine, Chicago, 1973.
- [Dun84] J. Duncan. Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113:501–517, 1984.
- [EAT92] M. J. Egenhofer and K. K. Al-Taha. Reasoning about gradual changes of topological relationships. In A.U. Frank, I. Campari, and U. Formentini, editors, *Theories and Methods of Spatio-temporal Reasoning in Geographic Space*, volume 639 of Lecture Notes in Computer Science, pages 196–219. Springer, Berlin, Heidelberg, 1992.
- [Edm93] P. G. Edmonds. A computational model of collaboration on reference in direction-giving dialogues. Technical Report CSRI-289, Computer Systems Research Institute, University of Toronto, 1993.
- [Edw54] W. Edwards. The theory of decision making. *Psychological Bulletin*, 51:380–417, 1954.
- [EE74] B. A. Eriksen and C. W. Eriksen. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, 16:143–149, 1974.
- [Ege91] M. J. Egenhofer. Reasoning about binary topological relations. In O. Günther and H.-J. Schek, editors, *Advances in Spatial Databases*, pages 144–160. Springer, Berlin, Heidelberg, 1991.
- [Ehr85] V. Ehrich. Zur linguistik und psycholinguistik der sekundären raumdeixis. In H. Schweizer, editor, *Sprache und Raum*, pages 130–161. Metzler, Stuttgart, 1985.
- [EL82] R. J. Elliott and M. E. Lesk. Route finding in street maps by computers and people. In *Proc. of AAAI-82*, pages 258–261, Pittsburgh, PA, 1982.
- [EM95] M. J. Egenhofer and D. M. Mark. Naive geography. In A. U. Frank and W. Kuhn, editors, *Spatial Information Theory: A Theoretical Basis for GIS*, pages 1–15. Springer, Berlin, Heidelberg, 1995.

- [Enk93] W. Enkelmann. Konzeption und nutzung von wissensrepräsentationen bei video-gestützten assistenzsystemen für die straßenfahrzeugführung. In O. Herzog, T. Christaller, and D. Schütt, editors, *Grundlagen und Anwendungen der Künstlichen Intelligenz (17. Fachtagung)*, pages 74–89. Springer, Berlin, Heidelberg, 1993.
- [ES86] C. W. Eriksen and J. D. St. James. Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics*, 40:225–240, 1986.
- [EZ83] J. Engelkamp and H. D. Zimmer. *Dynamic Aspects of Language Processing: Focus and Presupposition*. W. J. M. Levelt (ed.): Language and Communication 16. Springer, Berlin, Heidelberg, 1983.
- [EZ94] J. Engelkamp and H. Zimmer. *The Human Memory: a multimodal approach*. Hogrefe & Huber, 1994.
- [Fau85] G. Fauconnier. *Mental Spaces*. MIT Press, Cambridge, MA, 1985.
- [FB94] Ch. Freksa and Th. Barkowsky. On the relation between spatial concepts and geographic objects. Technical Report 43, Graduiertenkolleg *Kognitionswissenschaft*, Hamburg, 1994.
- [Fel85] J. A. Feldman. Four frames suffice: A provisional model of vision and space. *Behavioral and Brain Sciences*, 8:265–289, 1985.
- [Fel87] J. A. Feldman. A functional model of vision and space. In M.A. Arbib and A.P. Hanson, editors, *Vision, Brain, and Cooperative Computation*, pages 531–562. Bradford Book, MIT Press, Cambridge, MA, 1987.
- [FHLvB87] J. E. Fenstad, P.-K. Halvorsen, T. Langholm, and J. F. A. K. van Benthem. *Situations, Language and Logic*. Reidel, Dordrecht, Boston, 1987.
- [FHW93] FHWA. *IVHS Architecture Mission Definition*. Washington, DC, 1993.
- [Fil71] C. J. Fillmore. Toward a theory of deixis. Mimeographed, 1971. Paper read at the Pacific Conference on Contrastive Linguistics and Language Universals, University of Hawaii.
- [Fil75] C. J. Fillmore. *Santa Cruz Lectures on Deixis*. Indiana University Linguistic Club, Bloomington, IN, 1975. only reproduced, first published in 1971.
- [Fil82] C. J. Fillmore. Towards a descriptive framework for spatial deixis. In R.J. Jarvella and W. Klein, editors, *Speech, Place, and Action*, pages 31–59. Wiley, Chichester, 1982.
- [Fin96] W. Finkler. *Automatische Selbstkorrektur bei der inkrementellen Generierung gesprochener Sprache unter Realzeitbedingungen: Ein empirisch-simulierter Ansatz unter Verwendung eines Begründungsverwaltungssystems*. PhD thesis, Fachbereich Informatik, Universität des Saarlandes, Saarbrücken, 1996. In Vorbereitung.
- [FKR84] M. Fürsinn, M. Khenkhar, and B. Ruschkowski. Geosys — ein frage-antwortsystem mit räumlichem vorstellungsvermögen. In C.-R. Rollinger, editor, *Probleme des (Text-) Verstehens, Ansätze der künstlichen Intelligenz*, pages 172–184. Niemeyer, Tübingen, 1984.

- [FL90] A. D. Friederici and W. J. M. Levelt. Spatial reference in weightlessness: Perceptual factors and mental representations. *Perception & Psychophysics*, 47:253–266, 1990.
- [FM89] S. Feiner and K. McKeown. Coordinating text and graphics in explanation generation. In *Speech and Natural Language Workshop*. DARPA, 1989.
- [FO91] P. F. Fisher and T. M. Orf. An investigation of the meaning of near and close on a university campus. *Environment and Behavior*, 15:23–35, 1991.
- [FR93] Ch. Freksa and R. Röhrig. Dimensions of qualitative spatial reasoning. In *Qualitative Reasoning and Decision Technologies*, Barcelona, 1993. QUARDET'93.
- [Fra91] A. U. Frank. Qualitative spatial reasoning with cardinal directions. In H. Kaindl, editor, *7. Österreichische Artificial-Intelligence-Tagung*, pages 157–167. Springer, Berlin, Heidelberg, 1991.
- [Fre86a] G. Frege. Begriffsschrift, eine der arithmetischen nachgebildete formelsprache des reinen denkens (gekürzter nachdruck). In K. Berka and L. Kreiser, editors, *Logik-Texte: Kommentierte Auswahl zur Geschichte der Modernen Logik (vierte Auflage)*, pages 82–107. Akademie-Verlag, Berlin, 1986.
- [Fre86b] G. Frege. über sinn und bedeutung. In K. Berka and L. Kreiser, editors, *Logik-Texte: Kommentierte Auswahl zur Geschichte der Modernen Logik (vierte Auflage)*, pages 423–441. Akademie-Verlag, Berlin, 1986.
- [FRH92] W. Fastenmeier, G. Reichart, and R. Haller. Welche information brauchen fahrer wirklich? In *Das Mensch-Maschine-System im Verkehr*, pages 15–33. VDI Verlag, Düsseldorf, 1992.
- [FS86] R. A. Finke and R. N. Shepard. Visual functions of mental imagery. In K.R. Boff, L. Kaufman, and J. P. Thomas, editors, *Handbook of Perception and Human Performance*, volume II: Cognitive Processes and Performance, pages 37/1–37/55. Wiley, Chichester, 1986.
- [FS92] W. Finkler and A. Schauder. Effects of incremental output on incremental natural language generation. In *Proc. of the 10<sup>th</sup> ECAI*, pages 505–507, Vienna, 1992.
- [FTC92] N. Franklin, B. Tversky, and V. Coon. Switching points of view in spatial mental models. *Memory and Cognition*, 20:507–518, 1992.
- [Fur81] T. Furugori. Computing a map from michi-annai-bun or written directions. In *Proc. of the 7<sup>th</sup> IJCAI*, pages 426–428, Vancouver, Canada, 1981.
- [Gal90] C. Gallistel. *The Organization of Learning*. Cambridge, MA: MIT Press, 1990.
- [Gap94] K.-P. Gapp. On the basic meanings of spatial relations: Computation and evaluation in 3d space. In *Proc. of the 12<sup>th</sup> AAAI*, pages 1393–1398, Seattle, WA, 1994.
- [Gap95a] K.-P. Gapp. Angle, distance, shape, and their relationship to projective relations. In *Proc. of the 17th. Conf. of the Cognitive Science Society*, Pittsburgh, 1995.
- [Gap95b] K.-P. Gapp. Object localization: Selection of optimal reference objects. In A. Frank and W. Kuhn, editors, *Spatial Information Theory*, Lecture Notes in Computer Science, 988, pages 519–536. Springer, Berlin, Heidelberg, 1995.

- [Gap96] K.-P. Gapp. *Ein Berechnungsmodell für räumliche Relationen für dreidimensionale Umgebungen*. PhD thesis, Universität des Saarlandes, 1996.
- [Gar89] A. Garnham. A unified theory of the meaning of some spatial relational terms. *Cognition*, 31:45–60, 1989.
- [GBL85] T. Gärling, A. Böök, and E. Lindberg. Adult's memory representation of their everyday physical environment. In R. Cohen, editor, *The development of cognition*, pages 141–184. Erlbaum, Hillsdale, NJ, 1985.
- [GG87] T. Gärling and R. G. Golledge. Environmental perception and cognition. In E. H. Zube and G. T. Moore, editors, *Advances in environment, behavior, and design*, volume 2, pages 203–236. Plenum, New York, 1987.
- [Gib79] J. J. Gibson. *The Ecological Approach to Visual Perception*. Mifflin, Boston, MA, 1979.
- [GKS89] S. Gopal, R. Klatzky, and T. Smith. Navigator: A psychologically based model of environmental learning through navigation. *Journal of Environmental Psychology*, 9:309–331, 1989.
- [Gla94] J. Glasgow. Spatial reasoning in indeterminate worlds. In *Proc. of the 12<sup>th</sup> AAAI*, pages 1399–1404, Seattle, WA, 1994.
- [GLK<sup>+</sup>91] R. G. Golledge, J. M. Loomis, R. L. Klatzky, A. Flury, and X. L. Yang. Designing a personal guidance system to aid navigation without sight: progress on the gis component. *International Journal of Geographical Information Systems*, 5(4):373–395, 1991.
- [GM95] J. Grabowski and G. A. Miller. Factors affecting production and comprehension of spatial prepositions: A comparison of German and American English. Technical Report 56, Arbeiten der Forschungsgruppe "Sprache und Kognition", Universität Mannheim, 1995.
- [Gö89] G. Görz. Verarbeitung natürlicher Sprache. In K. von Luck, editor, *Künstliche Intelligenz. 7. Frühjahrsschule, KIFS-89*, pages 22–51. Springer, Berlin, Heidelberg, 1989.
- [Gol78] R. G. Golledge. Environmental cues, cognitive mapping, spatial behavior. In D. Bürkle et al., editor, *Behavior-environment research methods*, pages 35–46. Institute for Environmental Studies, University of Wisconsin, 1978.
- [Gol87] R. G. Golledge. *D. Stokols and I. Altman (Ed.), Handbook of Environmental Psychology*, volume 5, chapter 5: Environmental Cognition, pages 131–174. Wiley, Chichester, 1987.
- [Gol91] R. G. Golledge. Cognition of physical and built environments. In T. Gärling and G. W. Evans, editors, *Environment, Cognition, and Action: An Integrated Approach*, pages 35–62. Oxford University Press, New York, 1991.
- [Gol93] R. G. Golledge. Geographical perspectives on spatial cognition. In T. Gärling and R. G. Golledge, editors, *Behavior and Environment: Psychological and Geographical Approaches*, number 96 in *Advances in Psychology*, pages 16–46. North-Holland, Amsterdam, 1993.

- [Gop88] S. Gopal. *A Computational Model of Spatial Navigation*. PhD thesis, University of California, Santa Barbara, 1988.
- [GR73] R. G. Golledge and G. Rushton. Multidimensional scaling: Review and geographic applications. Technical Report 10, AAG Technical Report, Washington, DC, 1973.
- [Gri75] H. P. Grice. Logic and conversation. In P. Cole and J. L. Morgan, editors, *Speech Acts*, pages 41–58. Academic Press, London, 1975.
- [GRR82] R. G. Golledge, J.N. Rayner, and V.L. Rivizzigno. Comparing objective and cognitive representations of environmental cues. In R. G. Golledge and J.N. Rayner, editors, *Proximity and preference*. University of Minnesota Press, Minneapolis, 1982.
- [Gry92] A. Gryl. Opérations cognitives mises en oeuvre dans la description d'itinéraires. Technical Report 92, CNRS/LIMSI, Orsay Cedex, 1992.
- [GS78] R. G. Golledge and A.N. Spector. Comprehending the urban environment: theory and practice. *Geographical Analysis*, 10:403–426, 1978.
- [GS86] B. J. Grosz and C. Sidner. Attention, intention and the structure of discourse. *Computational Linguistics*, 12(3):175–204, 1986.
- [GW95] J. Grabowski and P. Weiß. The prepositional inventory of languages: A factor the affects comprehension of spatial prepositions. Technical Report 57, Arbeiten der Forschungsgruppe "Sprache und Kognition", Universität Mannheim, 1995.
- [GZ73] R. G. Golledge and G. Zannaras. Cognitive approaches to the analysis of human spatial behavior. In W.H. Ittelson, editor, *Environmental Cognition*, pages 59–94. Seminar Press, New York, 1973.
- [Hab86] Ch. Habel. *Prinzipien der Referentialität*. Springer, Berlin, Heidelberg, 1986.
- [Hab87] Ch. Habel. Prozedurale aspekte der wegplanung und wegbeschreibung. LILOG-Report 17, IBM, Stuttgart, 1987.
- [Hal55] A. I. Hallowell. *Culture and Experience*. University of Pennsylvania Press, Philadelphia, 1955.
- [Han80] K.-J. Hanßmann. Sprachliche bildinterpretation für ein frage-antwort-system. Bericht 74, Fachbereich Informatik, Universität Hamburg, 1980.
- [Hår66] A. Hård. A new color atlas based on the natural colour system by hering-johansson. *Die Farbe*, 15:296–304, 1966.
- [Hav92] J. Haviland. Anchoring, iconicity and orientation in guugu yimithirr pointing gestures. Technical Report 8, Cognitive Anthropology Research Group, Nijmegen, The Netherlands, 1992.
- [HB92] J. E. Hummel and I. Biederman. Dynamic binding in a neural network for shape recognition. *Psychological Review*, 99:480–517, 1992.
- [HBG<sup>+</sup>96] G. Herzog, A. Blocher, K.-P. Gapp, E. Stopp, and W. Wahlster. Vitra: Verbalisierung visueller information. *Informatik Forschung und Entwicklung*, 11(1):12–19, 1996.
- [HBN87] Th. Herrmann, B. Bürkle, and H. Nirmaier. Zur hörerbezogenen raumreferenz: Hörerposition und lokalisationsaufwand. Bericht 12, Forschergruppe "Sprechen und Sprachverstehen im sozialen Kontext", Heidelberg/Mannheim, 1987.

- [HCR90] W. Hoepfner, M. Carstensen, and U. Rhein. Wegauskünfte: Die Interdependenz von Such- und Beschreibungsprozessen. In C. Freksa and C. Habel, editors, *Informatik Fachberichte 245*, pages 221–234. Springer, Berlin, Heidelberg, 1990.
- [Hee91] P. A. Heeman. A computational model of collaboration on referring expressions. Technical Report CSRI-251, Computer science research institute, University of Toronto, 1991.
- [Her83] Th. Herrmann. *Speech and situation: A psychological conception of situated speaking*. Springer, Berlin, Heidelberg, 1983.
- [Her86] A. Herskovits. *Language and Spatial Cognition. An Interdisciplinary Study of the Prepositions in English*. Cambridge University Press, Cambridge, London, 1986.
- [Her87] Th. Herrmann. Sprachpsychologische beiträge zur partnerbezogenheit des sprechens. Bericht 37, Forschergruppe Sprache und Kognition, Lehrstuhl Psychologie III, Universität Mannheim, 1987.
- [Her89] Th. Herrmann. Partnerbezogene objektlokalisierung - ein neues sprachpsychologisches forschungsthema. Bericht 25, Arbeiten der Forschergruppe "Sprechen und Sprachverstehen im sozialen Kontext", Heidelberg/Mannheim, 1989.
- [Her90] Th. Herrmann. Vor, hinter, rechts und links: das 6h-modell. psychologische studien zum sprachlichen lokalisieren. *Zeitschrift für Literaturwissenschaft und Linguistik*, 78:117–140, 1990.
- [Her91] D. Hernandez. Relative representation of spatial knowledge: The 2-d case. In D. M. Mark and A. U. Frank, editors, *Cognitive and Linguistic Aspects of Geographic Space*, pages 373–385. Kluwer, Dordrecht, 1991.
- [Her92] G. Herzog. Utilizing interval-based event representations for incremental high-level scene analysis. In *Proceedings of the 4th. International Workshop on Semantics of Time, Space, and Movement and Spatio-Temporal Reasoning*, pages 425–435, Toulouse, 1992. Groupe "Langue, Raisonnement, Calcul".
- [Her96] G. Herzog. *Inkrementelle Szenenfolgenanalyse für den natürlichsprachlichen Zugang zu visuellen Daten*. PhD thesis, Universität des Saarlandes, 1996. in Vorbereitung.
- [HFS91] K. Harbusch, W. Finkler, and A. Schauder. Incremental syntax generation with tree adjoining grammars. Research Report RR-91-25, DFKI GmbH, Kaiserslautern und Saarbrücken, Deutschland, 1991.
- [HG94] Th. Herrmann and J. Grabowski. *Sprechen: Psychologie der Sprachproduktion*. Spektrum, Akademischer Verlag, 1994.
- [HGH91] Th. Herrmann, R. Graf, and E. Helmecke. "rechts" und "links" unter variablen betrachtungswinkeln: Nicht-shepardsche rotationen. Technical Report 37, Arbeiten aus dem Sonderforschungsbereich 245 "Sprechen und Sprachverstehen im sozialen Kontext", Universität Mannheim, 1991.
- [HH92] P. A. Heeman and G. Hirst. Collaborating on referring expressions. Technical Report TR 435, Computer Science Dept., University of Rochester, NY, 1992.
- [HHD91] J. Huttenlocher, L. V. Hedges, and S. Duncan. Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, 98:352–376, 1991.

- [Hil82] C. Hill. Up/down, front/back, left/right. a contrastive study of hausa and english. In J. Weissenborn and W. Klein, editors, *Here and There. Cross-Linguistic Studies on Deixis and Demonstration*. Benjamins, Philadelphia, PA, 1982.
- [HJ85] S. Hirtle and J. Jonides. Evidence of hierarchies in cognitive maps. *Memory and Cognition*, 13(3):208–217, 1985.
- [HK91] K. Höök and J. Karlgren. Some principles for route descriptions derived from human advisers. Technical Report R91:06, Swedish Institute of Computer Science, Kista, SE, 1991.
- [HKDD92] Th. Herrmann, E. Killian, E. Dittrich, and P. Dreyer. Was- und wie-schemata beim erzählen. In H. P. Krings and G. Antos, editors, *Textproduktion. Neue Wege der Forschung*, pages 147–158. WVT Wissenschaftlicher Verlag Trier, Trier, 1992. Fokus Bd. 7.
- [HLP78] N. L. Hazen, J. L. Lockman, and H. L. Pick. The development of children's representations of large-scale environments. *Child Development*, 49:623–636, 1978.
- [HM71] R. Hart and G. R. Moore. The development of spatial cognition: A review. Place Perception Report 7, Departement of Geography, Clark University, 1971.
- [HM73] R. Hart and G. R. Moore. The development of spatial cognition: A review. In R. M. Downs and D. Stea, editors, *Image and Environment*, pages 246–288. Aldine, Chicago, 1973.
- [HM82] K.J. Holyoak and W. A. Mah. Cognitive reference points in judgments of symbolic magnitude. *Cognitive Psychology*, 14:328–352, 1982.
- [HM92] S. C. Hirtle and M. F. Mascolo. The heuristics of spatial cognition. In *Proc. of the 13th. annual conference of the cognitive science society*, pages 629–634, Hillsdale, NJ, 1992. Erlbaum.
- [HM94] P. W. Halligan and J. C. Marshall, editors. *Spatial Neglect: Position Papers on Theory and Practice*, volume 4(2) of *A Special Issue*. Erlbaum Assoc., Hillsdale, NJ, 1994.
- [HMP76] D. A. Hardwick, C. W. McIntyre, and H. L. Pick. The content and manipulation of cognitive maps in children and adults. *Monographs of the Society for Research in Child Development*, 41, 1976.
- [HMRS95] B. Hildebrandt, R. Moratz, G. Rickheit, and G. Sagerer. Integration von bild- und sprachverstehen in einer kognitiven architektur. *Kognitionswissenschaft*, 4:118–128, 1995.
- [How86] I. P. Howard. The perception of posture, selfmotion, and the visual vertical. In K.R. Boff, L. Kaufman, and J. P. Thomas, editors, *Handbook of Perception and Human Performance*, volume II: Cognitive Processes and Performance, pages 18/1–18/62. Wiley, Chichester, 1986.
- [HP88] Ch. Habel and S. Pribbenow. Gebietskonstituierende prozesse. LILOG-Report 18, IBM, Stuttgart, 1988.
- [HR88] G. Herzog and T. Rist. Simultane interpretation und natürlichsprachliche beschreibung zeitveränderlicher szenen: Das system soccer. Memo 25, Universität des Saarlandes, SFB 314 (VITRA), 1988.

- [HR91] O. Herzog and C.-R. Rollinger, editors. *Text Understanding in LILOG: Integrating Computational Linguistics and Artificial Intelligence e. Final Report on the IBM Germany LILOG-Project*. Springer, Berlin, Heidelberg, 1991.
- [HR94] G.W. Humphreys and M.J. Riddoch. Attention to within-object and between-object spatial representations: Multiple sites for visual selection. *Cognitive Neuropsychology*, 11:207–241, 1994.
- [HRA89] G. Herzog, T. Rist, and E. André. Sprache und raum: Natürlichsprachlicher zugang zu visuellen daten. In C. Freksa and C. Habel, editors, *Repräsentation und Verarbeitung räumlichen Wissens*, pages 207–220. Springer, Berlin, Heidelberg, 1989.
- [HS78] J. F. Herman and A. W. Siegel. The development of cognitive mapping of the large-scale environment. *Journal of Experimental Child Psychology*, 1978:389–406, 1978.
- [HS84] M. Hußmann and P. Scheffe. The design of swyss, a dialogue system for scene analysis. In L. Bolc, editor, *Natural Language Communication with Pictorial Information Systems*, pages 143–201. Hanser/McMillan, München, 1984.
- [HSA<sup>+</sup>89] G. Herzog, C.-K. Sung, E. André, W. Enkelmann, H.-H. Nagel, T. Rist, W. Wahlster, and G. Zimmermann. Incremental natural language description of dynamic imagery. In Ch. Freksa and W. Brauer, editors, *Wissensbasierte Systeme. 3. Internationaler GI-Kongreß*, pages 153–162. Springer, Berlin, Heidelberg, 1989.
- [HT66] I. P. Howard and W. B. Templeton, editors. *Human Spatial Orientation*. Wiley, Chichester, 1966.
- [HT95] W. G. Hayward and M. J. Tarr. Spatial language and spatial representation. *Cognition*, 55:39–84, 1995.
- [HU90] D. P. Huttenlocher and S. Ullman. Recognizing solid objects by alignment with an image. *International Journal of Computer Vision*, 5(2):195–212, 1990.
- [HW94] G. Herzog and P. Wazinski. Visual translator: Linking perceptions and natural language descriptions. *Artificial Intelligence Review*, 8(2-3):175–187, 1994.
- [Itt73] W. H. Ittelson. Environment perception and contemporary perceptual theory. In W. H. Ittelson, editor, *Environment and Cognition*, pages 1–19. Seminar, New York, 1973.
- [Jac72] R. Jackendoff. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, MA, 1972.
- [Jac83] R. Jackendoff. *Semantics and Cognition*. MIT Press, Cambridge, MA, 1983.
- [Jac87] R. Jackendoff. *Consciousness and the Computational Mind*. MIT Press, Cambridge, MA, 1987.
- [Jac90] R. Jackendoff. *Semantic Structures*. MIT Press, 1990.
- [JL83] P. N. Johnson-Laird. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge University Press, 1983.

- [Jos85] A.K. Joshi. Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions. In D.R. Dowty et al., editor, *Natural Language Parsing: Psychological, computational and theoretical perspectives*, pages 206–250. Cambridge University Press, 1985.
- [JW75] D. B. Judd and G. Wyszecki. *Color in Business, Science, and Industry*. Wiley, Chichester, 1975.
- [Kau89] I. Kaufmann. Direktionale präpositionen. In C. Habel, M. Herweg, and K. Rehkämper, editors, *Raumkonzepte in Verstehensprozessen: Interdisziplinäre Beiträge zu Sprache und Raum*, pages 128–149. Niemeyer, Tübingen, 1989.
- [Kay79] M. Kay. Functional grammar. In *Proc of the 5th Annual Meeting of the Berkeley Linguistic Society*, 1979.
- [Ker96] J. H. Kerstholt. *Dynamic Decision Making*. PhD thesis, Universität Amsterdam, 1996.
- [KF95] A. Kilger and W. Finkler. Incremental generation for real-time applications. Research Report RR-95-11, Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI GmbH), 1995.
- [KH81] D. Kahneman and A. Henik. Perceptual organization and attention. In M. Kubovy and J. R. Pomerantz, editors, *Perceptual organization*. Erlbaum, Hillsdale, NJ, 1981.
- [KH87] G. Kempen and E. Hoenkamp. An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11(2):201–258, 1987.
- [Kin77] W. Kintsch. *Memory and Cognition*. Wiley, Chichester, 2 edition, 1977.
- [KJ91] A. F. Kramer and A. Jacobson. Perceptual organization and focussed attention: The role of objects and proximity in visual processing. *Perception and Psychophysics*, 50:267–284, 1991.
- [KL88] B. J. Kuipers and T.S. Levitt. Navigation and mapping in large-scale spaces. *AI Magazine*, pages 25–43, 1988.
- [Kle79] W. Klein. Wegauskünfte. *LiLi*, 9(33):9–57, 1979.
- [Kle82] W. Klein. Local deixis in route directions. In R. J. Jarvella and W. Klein, editors, *Speech, Place, and Action*, pages 161–182. Wiley, Chichester, 1982.
- [Kle94] W. Klein. Keine känguruhs zur linken - über die variabilität von raumvorstellungen und ihren ausdruck in der sprache. In H.-J. Kornadt, J. Grabowski, and R. Mangold-Allwinn, editors, *Sprache und Kognition*, pages 163–182. Spektrum Akademischer Verlag, Heidelberg, 1994.
- [Kli71] F. Klix. *Information und Verhalten*. Verlag Hans Huber, Bern, 1971.
- [Koe90] J. J. Koenderink. *Solid Shape*. MIT Press, Cambridge, MA, 1990.
- [Kor85a] R. E. Korf. Depth-first iterative-deepening: An optimal admissible tree search. *Artificial Intelligence*, 27:97–109, 1985.
- [Kor85b] R. E. Korf. Iterative-deepening-a\*: An optimal admissible tree search. In *Proc. of the 9<sup>th</sup> IJCAI*, pages 1034–1036, Los Angeles, CA, 1985.

- [Kor90] R. E. Korf. Real-time heuristic search. *Artificial Intelligence*, 42:189–211, 1990.
- [Kos80] S. M. Kosslyn. *Image and Mind*. Harvard University Press, Cambridge, MA, London, 1980.
- [Kos83] S. M. Kosslyn. *Ghosts in the Mind's Machine*. Norton, New York, 1983.
- [Kos87] S. Kosslyn. Seeing and imaging in the cerebral hemispheres: a computational approach. *Psychological Review*, 94:148–175, 1987.
- [Kos94] S. M. Kosslyn. *Image and Brain: The Resolution of the Imagery Debate*. MIT Press, Cambridge, MA, 1994.
- [KPF74] S. Kosslyn, H.L. Pick jr., and G.R. Fariello. Cognitive maps in children and men. *Child Development*, 45:707–716, 1974.
- [Kru74] L. Kruse. *Räumliche Umwelt. Die Phänomenologie des räumlichen Verhaltens als Beitrag zu einer psychologischen Umwelttheorie*. Berlin, New York, de Gruyter, 1974.
- [KT84] D. Kahneman and A. Treisman. Changing views of attention and automaticity. In R. Parasuraman and D.R. Davis, editors, *Varieties of attention*. Academic Press, Orlando, FL, 1984.
- [KTG92] D. Kahneman, A. Treisman, and B. Gibbs. The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24:175–219, 1992.
- [Kui77] B. J. Kuipers. *Representing Knowledge of Large-Scale Space*. PhD thesis, MIT AI Lab, Cambridge, MA, 1977. TR-418.
- [Kui78] B. J. Kuipers. Modeling spatial knowledge. *Cognitive Science*, 2:129–153, 1978.
- [Kui94] B. J. Kuipers. An ontological hierarchy for spatial knowledge. *AAAI-94 Fall Symposium on Control of the Physical World by Intelligent Agents*, 1994.
- [KW89] A. Kobsa and W. Wahlster, editors. *User Models in Dialog Systems*. Springer, Berlin, Heidelberg, 1989.
- [Lak72] G. Lakoff. Hedges: A study in meaning criteria and the logic of fuzzy concepts. In P. M. Peranteau, J. N. Levi, and G. C. Phares, editors, *Papers from the 8<sup>th</sup> Regional Meeting of the Chicago Linguistics Society*. University of Chicago, Department of Linguistics., Chicago, IL, 1972.
- [Lak73] G. Lakoff. Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*, 2:458–508, 1973.
- [Lak87] G. Lakoff. *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago University Press, Chicago, IL, 1987.
- [Lap66] J. S. Lappin. *Attention in the identification of stimuli in complex visual displays*. PhD thesis, Univ. of Illinois, Urbana, IL, 1966.
- [LB89] D. LaBerge and V. Brown. Theory of attentional operations in shape identification. *Psychological Review*, 96:101–124, 1989.
- [LCS90] E. Lang, K.-U. Carstensen, and G. Simmons. Modelling spatial knowledge on a linguistic basis. IWBS Report 147, Wissenschaftliches Zentrum, Institut für Wissensbasierte Systeme, IBM, Stuttgart, 1990.

- [Lei87] D. Leiser. The changing relations between representations and cognitive structures in the development of a cognitive map. *New Ideas in Psychology*, 5:95–110, 1987.
- [Lev82] W. J. M. Levelt. Cognitive styles in the use of spatial direction terms. In R. J. Jarvella and W. Klein, editors, *Speech, Place, and Action*, pages 251–268. Wiley, Chichester, 1982.
- [Lev83] W. J. M. Levelt. Monitoring and self-repair in speech. *Cognition*, 14:41–104, 1983.
- [Lev84] W. J. M. Levelt. Some perceptual limitations on talking about space. In A. van Doorn, W. van der Grind, and J. Koenderink, editors, *Limits in perception*, pages 323–358. VNU Science Press, Utrecht, 1984.
- [Lev86] W. J. M. Levelt. Zur sprachlichen abbildung des raumes: Deiktische und intrinsische perspektive. In H.-G. Bosshardt, editor, *Perspektiven auf Sprache*, pages 187–211. de Gruyter, Berlin, New York, 1986.
- [Lev89] W. J. M. Levelt. *Speaking — From Intention to Articulation*. MIT Press, Cambridge, MA, 1989.
- [Lev92] S. C. Levinson. Language and cognition: The cognitive consequences of spatial description in guugu yimithirr. Technical Report 13, Cognitive Anthropology Research Group, Nijmegen, The Netherlands, 1992.
- [Lev96] S. C. Levinson. Frames of reference and molyneux’s question: Cross-linguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel, and M. Garrett, editors, *Space and language*. MIT Press, Cambridge, MA, 1996. in press.
- [LFMM94] H. Lobin, S. Förster, J.-T. Milde, and G. Müller. Steps towards the realization of behavior-based agents. Report 94/3, SFB 360 - Situierte Kommunikatoren, Bielefeld, 1994.
- [LG83] E. Lindberg and T. Gärling. Acquisition of different types locational information in cognitive maps: automatic or effortful processing. *Psychological Research*, 45:19–38, 1983.
- [LGK95] J. M. Loomis, R. G. Golledge, and R. L. Klatzky. Personal guidance system for blind persons. In *Proc. of the Conf. on Orientation and Navigation Systems for Blind Persons*, Hatfield, England, 1995.
- [LGS81] B. Landau, H. Gleitman, and E. Spelke. Spatial knowledge and geometric representation in a child blind from birth. *Science*, 213:1275–1278, 1981.
- [LH87] R. Lloyd and C. Heivly. Systematic distortions in urban cognitive maps. *Annals of the Association of American Geographers*, 77:191–207, 1987.
- [LJ80] G. Lakoff and M. Johnson. *Metaphors we live by*. Chicago University Press, Chicago, IL, 1980.
- [LJ93] B. Landau and R. Jackendoff. “what” and “where” in spatial language and spatial cognition. *Behavioral and Brain Sciences*, 16:217–265, 1993.
- [LK79] G. R. Loftus and H. J. Kallman. Encoding and use of detail information in picture recognition. *Journal of Experimental Psychology: Human Learning and Memory*, 5:197–211, 1979.

- [LL75] C. Linde and W. Labov. Spatial network as a site for the study of language and thought. *Language*, 51:924–939, 1975.
- [Llo89] R. Lloyd. Cognitive maps: encoding and decoding information. *Annals of the Association of the American Geographers*, 79:101–124, 1989.
- [Log95] G. D. Logan. Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology*, 28:103–174, 1995.
- [Low87] D. G. Lowe. Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, 31:355–395, 1987.
- [LPN81] L. Liben, A. Patterson, and N. Newcombe, editors. *Spatial representation and behavior across the life span*. Academic Press, NY, 1981.
- [LSG84] B. Landau, E. Spelke, and H. Gleitman. Spatial knowledge in a young blind child. *Cognition*, 16:225–260, 1984.
- [LWF85] D. N. Levine, J. Warach, and M. Farah. Two visual systems in mental imagery: dissociation of 'what' and 'where' in imagery disorders due to bilateral posterior cerebral lesions. *Neurology*, 35:1010–1019, 1985.
- [Lyn60] K. Lynch. *The Image of the City*. MIT Press, Cambridge, MA, 1960.
- [LZ89] D. Leiser and A. Zilbershatz. The traveller: A computational model of spatial network learning. *Environmental and Behaviour*, 21(4):435–463, 1989.
- [Maa93] W. Maaß. A cognitive model for the process of multimodal, incremental route description. In A. U. Frank and I. Campari, editors, *Spatial Information Theory: A Theoretical Basis for GIS*, Lecture Notes in Computer Science, 716, pages 1–13. Springer, Berlin, Heidelberg, 1993.
- [Maa94] W. Maaß. From visual perception to multimodal communication: Incremental route descriptions. *Artificial Intelligence Review Journal*, 8(5/6), December 1994. Special Volume on Integration of Natural Language and Vision Processing.
- [Mac37] D. L. MacAdam. Projective transformations of icv color specifications. *J. of Opt. Soc. Am.*, 27:294–299, 1937.
- [Mai90] C. Maienborn. Position und bewegung: Zur semantik lokaler verben. LILOG-Report 138, IBM, Stuttgart, 1990.
- [Mak81] R. H. Maki. Categorization and distance effects with spatial linear orders. *Journal of Experimental Psychology: Human Learning and Memory*, 7:15–32, 1981.
- [Man83] J. M. Mandler. Representation. In P. Mussen, editor, *Handbook of child psychology*, volume III (4th ed.), pages 420–494. Wiley, Chichester, 1983.
- [Man86] R. Mangold. *Sensorische Faktoren beim Verstehen überspezifizierter Objektbenennungen*. Lang, Frankfurt/Main, 1986.
- [Mar82] D. Marr. *Vision: a computational investigation into the human representation and processing of visual information*. Freeman, San Francisco, 1982.
- [Mar89] D. M. Mark. A conceptual model for vehicle navigation systems. In *Vehicle Navigation & Information Systems Conference*, pages 448–453, Toronto, 1989.

- [May92] M. T. Maybury. Spatial models for language processing. In *Workshop at the ECAI-92*, pages 47–53, 1992.
- [MB83] I. Moar and G. H. Bower. Inconsistency in spatial knowledge. *Memory and Cognition*, 11:107–113, 1983.
- [MBP95] W. Maaß, J. Baus, and J. Paul. Visual grounding of route descriptions in dynamic environments. In *AAAI Fall Symposium on Computational Models for Integrating Language and Vision*, MIT, Cambridge, MA, 1995. AAAI.
- [MC82] I. Moar and L. Carlton. Memory for routes. *Quarterly Journal of Experimental Psychology*, 34A:381–394, 1982.
- [McK94a] P. McKeivitt, editor. *Integration of Natural Language and Vision Processing*. AAAI-94 Workshop. Seattle, WA, 1994.
- [McK94b] P. McKeivitt, editor. *Special Volume on the Integration of Natural Language and Vision Processing*, volume 8 of *Artificial Intelligence Review Journal*. Kluwer, Dordrecht, 1994.
- [McN86] T. P. McNamara. Mental representations of spatial relations. *Cognitive Psychology*, 18:87–121, 1986.
- [McN92] T. P. McNamara. Spatial representation. *Geoforum*, 23:139–150, 1992.
- [MD84] D. McDermott and E. Davis. Planning and executing routes through uncertain territory. *Artificial Intelligence*, 22:107–156, 1984.
- [ME85] H. Moravec and A. Elfes. High resolution maps from wide angle sonar. In *International Conference on Robotics and Automation*, pages 116–121. IEEE, 1985.
- [Meh65] A. Mehrabian. *Räume des Alltags oder Wie die Umwelt unser Verhalten bestimmt*. Suhrkamp, Frankfurt, 1965.
- [MHH92a] T. P. McNamara, J. Halpin, and J. Hardy. The representation and integration in memory of spatial and nonspatial information. *Memory and Cognition*, 20(5):519–532, 1992.
- [MHH92b] T. P. McNamara, J. Halpin, and J. Hardy. Spatial and temporal contributions to the structure of spatial memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 18:555–564, 1992.
- [MJ76] S. Milgram and D. Jodelet. Psychological maps of paris. In H. Proshansky, W. Ittelson, and L. Rivlin, editors, *Readings in environmental psychology*, pages 104–124. Holt, Rinehart & Winston, New York, 1976.
- [MJL76] G. A. Miller and P. N. Johnson-Laird. *Language and Perception*. Cambridge University Press, Cambridge, London, 1976.
- [MMP<sup>+</sup>88] J. Meier, D. Metzger, T. Polzin, P. Ruhrberg, H. Rutz, and M. Vollmer. Generierung von Wegbeschreibungen. KoLiBri Arbeitsbericht 9, Fakultät für Linguistik und Literaturwissenschaft, Universität Bielefeld, 1988.
- [MN78] D. Marr and H. K. Nishihara. Representation and recognition of the spatial organization of three-dimensional shapes. In *Proc. Royal Society of London B*, pages 269–294, 1978.

- [MN87] J. H. R. Maunsell and W. T. Newsome. Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, 10:363–401, 1987.
- [Mon74] R. Montague. *Formal Philosophy, Selected Papers of Richard Montague*. New Haven and London, 1974. Ed. and with an introduction by R.H. Thomason.
- [Mon93] D. R. Montello. Scale and multiple psychologies of space. In A. U. Frank and I. Campari, editors, *Spatial Information Theory: A Theoretical Basis for GIS*, number 716 in Lecture Notes in Computer Science, pages 312–321. Springer, Berlin, Heidelberg, 1993.
- [MS79] G. McCalla and P. Schneider. The execution of plans in an independent dynamic microworld. In *Proc. of the 6<sup>th</sup> IJCAI*, pages 553–555, Tokio, 1979.
- [Mül88] S. Müller. CITYGUIDE: Ein System zur Wegplanung und Wegbeschreibung. Master's thesis, Fachbereich Informatik der Universität des Saarlandes, 1988.
- [Mül95] J. Müller. Geodisplay: Ein flexibles 3d-visualisierungswerkzeug fuer lisp. Master's thesis, Fachbereich Informatik der Universität des Saarlandes, Saarbrücken, Deutschland, 1995.
- [MUM83] M. Mishkin, L. Ungerleider, and K. Macko. Object vision and spatial vision: two cortical pathways. *Trends Neuroscience*, 6:414–417, 1983.
- [MWH93] W. Maaß, P. Wazinski, and G. Herzog. Vitra guide: Multimodal route descriptions for computer assisted vehicle navigation. In *Proc. of the Sixth International Conference on Industrial and Engineering Applications on Artificial Intelligence and Expert Systems*, Edinburgh, U.K., June 1-4,1993.
- [MZ92] J. L. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision*. MIT Press, Cambridge, MA, 1992.
- [Nad88] L. Nadel. Landmarks: Neurobiological perspectives. *British Journal of Developmental Psychology*, 6:383–385, 1988.
- [Nag88] H.-H. Nagel. From image sequences towards conceptual descriptions. *Image and Vision Computing*, 6(2):59–74, 1988.
- [NCG93] Progress report - caltrans agreement 65t115. Technical report, National Center for Geographic Information and Analysis (NCGIA), UC Santa Barbara, 1993.
- [Nei76] U. Neisser. *Cognition and Reality*. Freeman, San Francisco, 1976.
- [Neu86] B. Neumann. Optical flow. In N. I. Badler and J. K. Tsotsos, editors, *Motion: Representation and Perception*, pages 109–120. Elsevier, Amsterdam, 1986.
- [Neu89] B. Neumann. Natural language description of time-varying scenes. In D. L. Waltz, editor, *Semantic Structures: Advances in Natural Language Processing*, pages 167–207. Erlbaum, Hillsdale, NJ, 1989.
- [New88] N. Newcombe. The paradox of proximity in early spatial representation. *British Journal of Developmental Psychology*, 6:376–378, 1988.
- [NJ94] A. Ndiaye and A. Jameson. Supporting flexibility and transmutability: Multi-agent processing and role-switching in a pragmatically oriented dialog system. In Ph. Jorrand and V. Sgurev, editors, *Proceedings of the Sixth International Conference on Artificial Intelligence*, pages 381–390, Singapore, 1994.

- [NL82] N. Newcombe and L. S. Liben. Barrier effects in the cognitive maps of children and adults. *Journal of Experimental Child Psychology*, 34:46–58, 1982.
- [NN83a] B. Neumann and H.-J. Novak. Event models for recognition and natural language description of events in real-world image sequences. In *Proc. of the 8<sup>th</sup> IJCAI*, pages 724–726, Karlsruhe, FRG, 1983.
- [NN83b] B. Neumann and H.-J. Novak. Natural-language oriented event models for image sequences interpretation: The issues. Technical Note 34, CSRG, University of Toronto, 1983.
- [Nun79] G. Nunberg. The non-uniqueness of semantic solutions. *Linguistics and Philosophy*, 3:143–184, 1979.
- [OAV93] B. A. Olshausen, C. H. Anderson, and D. C. Van Essen. A multiscale routing circuit for forming size- and position-invariant object representations. *The Journal of Computational Neuroscience*, 2:45–62, 1993.
- [OB83] D. R. Olson and E. Bialystok. *Spatial Cognition: The Structure and Development of Mental Representations of Spatial Relations*. Erlbaum, Hillsdale, NJ, 1983.
- [OK95] B. A. Olshausen and C. Koch. Selective visual attention. In M. Arbib, editor, *The Handbook of Brain Theory and Neural Networks*. MIT Press, Cambridge, MA, 1995.
- [OMT94] P. Olivier, T. Maeda, and Jun-ichi T. Automatic depiction of spatial descriptions. In *Proc. of the 12<sup>th</sup> AAAI*, pages 1405–1410, Seattle, WA, 1994.
- [Pai70] J. Pailhous. *La Representation de l'Espace Urbain - L'exemple du Chauffeur de Taxi*. Presses Universitaires de France, 1970.
- [Pai71] A. Paivio. *Imagery and Verbal Processes*. Holt, Rinehart and Winston, New York, 1971.
- [Pau95] J. Paul. Moses-speak: Planbasierte generierung multimodaler inkrementeller wegbeschreibungen. Master's thesis, Fachbereich Informatik, Universität des Saarlandes, 1995. in Vorbereitung.
- [PB95] C. Pinhanez and A. F. Bobick. Scripts in machine understanding of image sequences. In *AAAI-95 Fall Symposium on "Computational Models for Integrating Language and Vision"*, pages 96–101, Cambridge, MA, 1995. AAAI.
- [PC84] M. I. Posner and Y. Cohen. Components of visual orienting. In H. Bouma and D. Bouwhuis, editors, *Attention and Performance X*. Erlbaum, Hillsdale, NJ, 1984.
- [PC90] T. Pattabhiraman and N. Cercone. Selection: Saliency, relevance and the coupling between domain-level tasks and text planning. In *Fifth International Workshop on Natural Language Generation*, pages 79–86, Dawson, PA, 1990.
- [PE88] D. V. Pullar and M. Egenhofer. Towards formal definitions of topological relations among spatial data handling. In *Third International Symposium on Spatial Data Handling*, pages 225–241, Sydney, Australia, 1988.
- [Ped93] E. Pederson. Geographic and manipulable space in two tamil linguistic systems. In A. U. Frank and I. Campari, editors, *Spatial Information Theory: A Theoretical Basis for GIS*, Lecture Notes in Computer Science, 716, pages 294–309. Springer, Berlin, Heidelberg, 1993.

- [PI67] J. Piaget and B. Inhelder. *The child's conception of space*. Norton, New York, 1967. Originally published in French, 1948.
- [Pia26] J. Piaget. *Judgement and Reasoning in the Child*. Harcourt Brace, 1926.
- [PIFC87] M. I. Posner, A.W. Inhoff, F. J. Friedrich, and A. Cohen. Isolating attentional systems: A cognitive-anatomical analysis. *Psychobiology*, 15:107–121, 1987.
- [Pin84] S. Pinker. Visual cognition: An introduction. *Cognition*, 18:1–63, 1984.
- [Pin92] M. Pinkal. Semantik. In G. Görz, editor, *AI*, pages 425–497. Addison-Wesley, 1992.
- [PIS60] J. Piaget, B. Inhelder, and A. Szemenska. *The Child's Conception of Geometry*. Basic Books, New York, 1960.
- [PM88] C. C. Presson and D. R. Montello. Points of reference in spatial cognition: Stalking the elusive landmark. *British Journal of Developmental Psychology*, 6:378–381, 1988.
- [PM94] C. C. Presson and D. R. Montello. Updating after rotational and translational body movements: coordinate structure of perspective space. *Perception*, 23:1447–1455, 1994.
- [PMS88] H. L. Pick, D. R. Montello, and S. C. Somerville. Landmarks and the coordination and integration of spatial information. *British Journal of Developmental Psychology*, 6:372–375, 1988.
- [Pob91] R. Pobel. *Objektpräsentation und Objektbenennung: situative Einflüsse auf die Wortwahl beim Benennen von Gegenständen*. Roderer, Regensburg, 1991.
- [Pre82] C. C. Presson. The development of map reading skills. *Child development*, 53:196–199, 1982.
- [Pre87] C. C. Presson. The development of spatial cognition: Secondary uses of spatial information. In N. Eisenberg, editor, *Contemporary Topics in Developmental Psychology*. Wiley, Chichester, 1987.
- [Pri91] S. Pribbenow. *Zur Verarbeitung von Lokalisierungsdrücken in einem hybriden System*. PhD thesis, Fachbereich Informatik, Universität Hamburg, 1991.
- [Pri93] S. Pribbenow. *Räumliche Konzepte in Wissens- und Sprachverarbeitung*. Deutscher Universitätsverlag, 1993.
- [Pyl89] Z. Pylyshyn. The role of location indexes in spatial perception: A sketch of the first spatial-index model. *Cognition*, 32:65–97, 1989.
- [PYR79] H. L. Pick, A. Yonas, and J. J. Rieser. Spatial reference systems in perceptual development. In N. Bornstein and W. Kessen, editors, *Psychological Development from Infancy: Image to Intention*. Erlbaum, Hillsdale, NJ, 1979.
- [R94] R. Röhrig. A theory for spatial reasoning based on order relations. In *Proc. of the 12<sup>th</sup> AAAI*, pages 1418–1423, Seattle, WA, 1994.
- [RCC92] D. A. Randall, Z. Cui, and A. G. Cohn. A spatial logic based on regions and connection. In *Proc. of the 3rd Int. Conf. on Knowledge Representations and Reasoning*, pages 165–176, San Mateo, 1992. Morgan Kaufmann.

- [RD92] E. Reiter and R. Dale. A fast algorithm for the generation of referring expressions. In *Proceedings of the 14th. International Conference on Computational Linguistics*, pages 232–238. COLING-92, 1992.
- [RHA92] Rhapit: Rhein/main project for integrated traffic management. Technical report, DRIVE, European Community, 1992.
- [Ric82] G. Richardson. *Spatial Cognition*. PhD thesis, University of California, Santa Barbara, 1982.
- [Rie89] J. J. Rieser. Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15:1157–1165, 1989.
- [RM81] R. Ratcliff and G. McKoon. Does activation really spread? *Psychological Review*, 88:454–462, 1981.
- [Rob63] L. G. Roberts. *Machine perception of three-dimensional solids*. PhD thesis, MIT Lincoln Laboratory, 1963.
- [Roe59] C. O. Roelofs. Considerations on the visual egocentre. *Acta psychologica*, 16:226–234,274,277,278, 1959.
- [Ros73] E. H. Rosch. Natural categories. *Cognitive Psychology*, 4:328–350, 1973.
- [RR90] P. Ruhrberg and H. Rutz. Räumliches wissen und semantik im kontext der generierung von wegbeschreibungen. In C. Freksa and C. Habel, editors, *Informatik Fachberichte 245*, pages 235–249. Springer, Berlin, Heidelberg, 1990.
- [RS87] A. Rau and M. Schweitzer. Untersuchungen zur Verwendung scriptartiger Wissensrepräsentationen und Story Grammars bei der Generierung von Wegbeschreibungen. Paper B 158, L.A.U.D., Duisburg, 1987.
- [RS88] G. Retz-Schmidt. Various views on spatial prepositions. *AI Magazine*, 9(2):95–105, 1988.
- [RS92] G. Retz-Schmidt. *Die Interpretation des Verhaltens mehrerer Akteure in Szenenfolgen*. Springer, Berlin, Heidelberg, 1992.
- [Rum77] D. E. Rumelhart. Toward an interactive model of reading. In S. Dornic, editor, *Attention and Performance IV*. Erlbaum, Hillsdale, NJ, 1977.
- [Rup96] U. Rupp. Grator - räumliches schliessen mit gradierten topologischen relationen über punktmengen. Master's thesis, Fachbereich Informatik, Universität des Saarlandes, 1996.
- [RW91] S. Russell and E. Wefald. *Do the Right Thing: Studies in Limited Rationality*. MIT Press, Cambridge, MA, 1991.
- [SA77] R. C. Schank and R. P. Abelson. *Scripts, Plans, Goals and Understanding*. Erlbaum, Hillsdale, NJ, 1977.
- [Sad88] E. Sadalla. Landmarks in memory. *British Journal of Developmental Psychology*, 6:386–388, 1988.
- [Sai84] G. Saile. *Sprache und Handlung*. Vieweg, Braunschweig, 1984.

- [SB55] R. Swanson and A. Benton. Some aspects of the genetic development of right-left discrimination. *Child Developm.*, 26:123–133, 1955.
- [SBS80] E. Sadalla, W. Burroughs, and L. Staplin. Reference points in spatiala cognition. *Journal of Experimental Psychology: Human Learning and Memory*, 5:515–528, 1980.
- [SBSZ87] J. R. J. Schirra, G. Bosch, C.-K. Sung, and G. Zimmermann. From image sequences to natural language: A first step towards automatic perception and description of motions. *Applied Artificial Intelligence*, 1:287–305, 1987.
- [SC78] A. Stevens and P. Coupe. Distortions in judged spatial relations. *Cognitive Psychology*, 10:422–437, 1978.
- [Sch87] E.A. Schegloff. Between macor and micro: Contexts and other connections. In J. Alexander/B. Giesen/R. Munch/ N. Smelser, editor, *The Micro-Macro Link*. Univ. of Cal Press, Berkeley, 1987.
- [Sch92] D. Schmauks. Wegmarkierungen, in: K. robering (hg.): Wegweiser zur bedeutung. *Arbeitspapiere zur Linguistik der TU Berlin*, 28:99–112, 1992. erscheint auch in: Zeitschrift für Semiotik.
- [Sch93] M. F. Schober. Spatial perspective-taking in conversation. *Cognition*, 47:1–24, 1993.
- [Sch94] J. R. J. Schirra. *Bildbeschreibung als Verbindung von visuellem und sprachlichem Raum – Eine interdisziplinäre Untersuchung von Bildvorstellungen in einem Hörermodell*. PhD thesis, Fachbereich Informatik, Universität des Saarlandes, Saarbrücken, 1994.
- [Sch95a] Ch. Schlieder. Reasoning about ordering. In A. U. Frank and W. Kuhn, editors, *Spatial Information Theory: A Theoretical Basis for GIS*, pages 341–350. Springer, Berlin, Heidelberg, 1995.
- [Sch95b] M. F. Schober. Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about locations? *Discourse Processes*, 20:219–247, 1995.
- [Sea69] J. R. Searle. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge, London, 1969.
- [Sen85] K. Sennholz. *Grundzüge der Deixis*. Brockmeyer, Bochum, 1985.
- [SGH<sup>+</sup>94] E. Stopp, K.-P. Gapp, G. Herzog, T. Laengle, and T. C. Lueth. Utilizing spatial relations for natural language access to an autonomous mobile robot. In B. Nebel and L. Dreschler-Fischer, editors, *KI-94: Advances in Artificial Intelligence*, pages 39–50. Springer, Berlin, Heidelberg, 1994.
- [SH84] R. N. Shepard and S. Hurwitz. Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition*, 18:161–193, 1984.
- [She62] F. Shemyakin. *On the psychology of space representations*. Uchenye Zapiski Gos. In-ta Psikhologii (Mokau), 1962. 1, n.p. Cited in Shemyakin, Orientation in space, in B. Ananyev (eds.) *Psychological Science in the U.S.S.R.*, Vol. 1, Washington: Office of Technical Services, 1962, 186-255.
- [Sho87] M. J. Sholl. Cognitive maps as orienting schemata. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13:615–628, 1987.

- [Sho92] M. J. Sholl. Landmarks, places, environments: Multiple mind-brain systems for spatial orientation. *Geoforum*, 23(2):151–164, 1992.
- [Shu84] V. Shute. *Characteristics of cognitive cartography*. PhD thesis, University of California, Santa Barbara, 1984.
- [SK96] M. Streit and A. Krüger. Eine agentenorientierte architektur für multimedial benutzerschnittstellen. In W. Wahlster, editor, *ONLINE96*, pages C622.01–C622.18, Velbert, 1996. Online GmbH.
- [SL88] E. Stern and D. Leiser. Levels of spatial knowledge and urban travel modelling. *Geographical Analysis*, 20(2):140–155, 1988.
- [SL96] E. Stopp and T. Laengle. Natürlichsprachliche instruktionen an einen autonomen serviceroboter. In R. Dillmann, U. Rembold, and T. Lth, editors, *Autonome Mobile Systeme*, pages 299–308, Berlin, Heidelberg, 1996. Springer. Tagungsband der “Autonome Mobile Systeme“(AMS ’95).
- [SM89] E. Sadalla and D. R. Montello. Remembering changes in direction. *Environment and Behavior*, 21:346–363, 1989.
- [Som69] R. Sommer. *Personal space. The behavioral basis of design*. Prentice Hall, Englewood Cliffs, 1969.
- [SS80] E. Sadalla and L. Staplin. The perceptions of traversed distance cognition. *Environment and Behavior*, 12:183–193, 1980.
- [SS93] J. R. J. Schirra and E. Stopp. Antlima: A listener model with mental images. In *Proc. of the 13<sup>th</sup> IJCAI*, pages 175–180, Chambéry, France, 1993.
- [Ste69] D. Stea. The measurement of mental maps: an experimental model for studying conceptual spaces. In K. R. Cox and R. G. Golledge, editors, *Behavioral Problems in Geography: a Symposium*, pages 228–253, Evanston, IL, 1969. Northwestern University Press.
- [Ste78] A. Stevens. *The role of inference and internal structure in representation of spatial information*. PhD thesis, Department of Psychology, University of California, San Diego, 1978.
- [Sv96] B. A. C. Saunders and J. van Brakel. Are there non-trivial constraints on colour categorization? *Behavioral and Brain Sciences*, 1996. under peer review.
- [SVW85] L. A. Streeter, D. Vitello, and S. A. Wonsiewicz. How to tell people where to go: Comparing navigational aids. *International Journal of Man-Machine Studies*, 22:549–562, 1985.
- [SW75] A. W. Siegel and S. H. White. The development of spatial representation of large-scale environments. In W. Reese, editor, *Advances in Child Development and Behaviour*, volume 10, pages 9–55. Academic Press, New York, 1975.
- [Tal83] L. Talmy. How language structures space. In H. L. Pick and L. Acredolo, editors, *Spatial Orientation: Theory, Research and Application*, pages 225–282. Plenum, New York, London, 1983.

- [Tal96] L. Talmy. Fictive motion in language and "ception". In P. Bloom, M. Peterson, and M. Garrett, editors, *Language and Space*. MIT Press, Cambridge, MA, 1996. Forthcoming.
- [TG80] A. Treisman and G. Gelade. A feature integration theory of attention. *Cognitive Psychology*, 12:97–136, 1980.
- [THKS91] C. Thorpe, M. H. Hebert, T. Kanade, and S. A. Shafer. Vision and navigation for the carnegie-mellon navlab. In S. S. Iyengar and A. Elfes, editors, *Autonomous Mobile Robots: Control, Planning, and Architecture (Vo l. 2)*, pages 401–412. IEEE Computer Society Press, Los Alamitos, CA, 1991.
- [Tho77] H. Thompson. Strategy and tactics in language production. In *Papers from the 13th. Regional Meeting of the Chicago Linguistics Society*, pages April 14–16. W.A. Beach and S.E. Fox and S. Philosoph, Chicago, 1977.
- [Tho81] P. Thorndyke. Distance estimations from cognitive maps. *Cognitive Psychology*, 13:526–550, 1981.
- [THR82] P.W. Thorndyke and B. Hayes-Roth. Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, 14:560–582, 1982.
- [TLS<sup>+</sup>95] W. Tack, K. Lorenz, J. Siekmann, H. Engelkamp, H. Zimmer, W. Wahlster, M. Kohlhase, A. Jameson, G. Smolka, M. Pinkal, and H. Uszkoreit. Antrag an die dfg auf einrichtung und förderung des sfb 1542: Ressourcenadaptive kognitive prozesse, 1995.
- [Tob70] W. R. Tobler. A computer movie simulating urban growth in the detroit region. *Economic Geography Supplement*, 46:234–240, 1970.
- [Tob76] W. R. Tobler. The geometry of mental maps. In R.G. Golledge and R. Rushton, editors, *Spatial choice and spatial behavior*, pages 69–81. Ohio State University Press, Columbus, 1976.
- [Tod62] M. Toda. The design of a fungus-eater: A model of human behavior in an unsophisticated environment. *Behavioral Science*, 7:164–183, 1962.
- [Tol48] E. C. Tolman. Cognitive maps in rats and men. *Psychological Review*, 55:189–208, 1948.
- [Tre88] A. Treisman. Features and objects: The fourteenth bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology*, 40a(2):201–237, 1988.
- [Tre93] A. Treisman. The perception of features and objects. In A. Baddeley and L. Weiskrantz, editors, *Attention: Selection, awareness and control. A tribute to Donald Broadbent*, pages 5–35. Clarendon Press, Oxford, 1993.
- [Tro13] C. C. Trowbridge. On fundamental methods of orientation and 'imagery maps'. *Science*, 38:888–897, 1913.
- [Tso85] J. K. Tsotsos. Knowledge organization and its role in representation and interpretation for time-varying data: the alven system. *Computational Intelligence*, 1:16–32, 1985.
- [Tso86] J. K. Tsotsos. The scope of research on motion: Sensations, perception, representation and generation. In N. I. Badler and J. K. Tsotsos, editors, *Motion: Representation and Perception*, pages 20–26. Elsevier, Amsterdam, 1986.

- [Tso90] J. K. Tsotsos. Analyzing vision at the complexity level. *Behavioral and Brain Sciences*, 13:423–469, 1990.
- [TT92] H. Taylor and B. Tversky. Descriptions and depictions of environments. *Memory and Cognition*, 20(5):483–496, 1992.
- [Tve81] B. Tversky. Distortions in memory for maps. *Cognitive Psychology*, 13:407–433, 1981.
- [Tve92] B. Tversky. Distortions in cognitive maps. *Geoforum*, 23:131–138, 1992.
- [Tve93] B. Tversky. Cognitive maps, cognitive collages, and spatial mental models. In A. U. Frank and I. Campari, editors, *Spatial Information Theory: A Theoretical Basis for GIS*, pages 14–24. Springer, Berlin, Heidelberg, 1993.
- [TVPE92] S. Timpf, G. S. Volta, D. W. Pollock, and M. J. Egenhofer. A conceptual model of wayfinding using multiple levels of abstraction. In A. U. Frank, I. Campari, and U. Formentini, editors, *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space: Proc. of the International Conference GIS*, pages 348–367. Springer, Berlin, Heidelberg, 1992.
- [UE82] V. Ullmer-Ehrich. The structure of living space descriptions. In R. J. Jarvella and W. Klein, editors, *Speech, Place, and Action*, pages 219–249. Wiley, Chichester, 1982.
- [Ull85] S. Ullman. Visual routines. In S. Pinker, editor, *Visual Cognition*, pages 97–159. MIT Press, Cambridge, MA, 1985.
- [Usz86] H. Uszkoreit. Categorical unification grammars. In *Proc. 11th. International Conference on Computational Linguistics*, pages 187–194, 1986.
- [van92] A. H. C. van der Heijden. *Selective attention in vision*. Routledge, London, New York, 1992.
- [VB90] S. Vere and T. Bickmore. A basic agent. *Computational Intelligence*, 6:41–60, 1990.
- [vC93] Ch. von Stutterheim and M. Carroll. Raumkonzepte in produktionsprozessen. *Kognitionswissenschaft*, 3:70–82, 1993.
- [Ver85] S. Vere. Temporal scope of assertions and window cutoff. In *Proc. of the 9<sup>th</sup> IJCAI*, pages 1055–1059, Los Angeles, CA, 1985.
- [vM83] D. van Essen and J. H. R. Maunsell. Hierarchical organization and functional streams in the visual cortex. *Trend in Neuroscience*, 6:370–375, 1983.
- [WAF<sup>+</sup>92] W. Wahlster, E. Andre, W. Finkler, W. Graf, H.-J. Profitlich, T. Rist, and A. Schauder. Wip: Integrating text and graphics design for adaptive information presentation. In R. Dale, E. Hovy, D. Rösner, and O. Stock, editors, *Aspects of Automated Natural Language Generation: Proc. of the 6th. International Workshop on Natural Language Generation*, pages 290–292. Springer, Berlin, Heidelberg, 1992.
- [Wah82] W. Wahlster. Natürlichsprachliche systeme. eine einföhrung in die sprachorientierte ki-forschung. In W. Bibel and J. Siekmann, editors, *Künstliche Intelligenz*, pages 203–283. Springer, Berlin, Heidelberg, 1982.

- [Wah91] W. Wahlster. Wie komme ich zum finanzamt? zur automatischen generierung sprachlicher beschreibungen aus visuellen daten. *D. Babel (Hrsg.): Künstliche Intelligenz, Zeitschrift Saarbrücker Hefte*, 65:33–42, 1991.
- [Wah94] W. Wahlster. Text and images. In R. A. Cole, J. Mariani, H. Uszkoreit, A. Zaenen, and V. Zue, editors, *Survey on Speech and Natural Language Technology*. Kluwer, Dordrecht, 1994.
- [WC94] I. Wachsmuth and Y. Cao. Interactive graphics design with situated agents. In W. Strasser and F. Wahl, editors, *Graphics and Robotics*. Springer, Berlin, Heidelberg, 1994.
- [Wil75] R. Wilton. Knowledge of spatial relations. a specification from the information used in making inferences. *Quarterly Journal of Experimental Psychology*, 31:133–146, 1975.
- [WJH78] W. Wahlster, A. Jameson, and W. Hoepfner. Glancing, referring and explaining in the dialogue system ham-rpm. *American Journal of Computational Linguistics, Microfiche 77:53–67*, 1978.
- [WR82] D. Wunderlich and R. Reinelt. How to get there from here. In R. J. Jarvella and W. Klein, editors, *Speech, Place, and Action*, pages 183–201. Wiley, Chichester, 1982.
- [WS82] G. Wyszecki and W. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Wiley, Chichester, 1982.
- [Wun78] D. Wunderlich. Wie analysiert man gespräche? beispiel wegauskünfte. *Linguistische Berichte*, 58:41–76, 1978.
- [Wun81] D. Wunderlich. Linguistic strategies. In F. Coulmas, editor, *A Festschrift for Native Speaker*, pages 279–296. Mouton, The Hague, 1981.
- [Wun82] D. Wunderlich. Sprache und raum. *Studium Linguistik*, 12/13:1–19, 37–59, 1982.
- [Wun86] D. Wunderlich. Raum und die struktur des lexikons. In H.-G. Bosshardt, editor, *Perspektiven auf Sprache*, pages 212–231. de Gruyter, Berlin, New York, 1986.
- [Yea88] W. K. Yeap. Towards a computational theory of cognitive maps. *Artificial Intelligence*, 34:297–360, 1988.
- [Zad65] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
- [Zek81] S. M. Zeki. The mapping of visual functions in the cerebral cortex. In Y. Katsuki, R. Norgren, and M. Sato, editors, *Brain mechanisms of sensation*. Wiley, Chichester, 1981.
- [Zim80] H. D. Zimmer. Feature und feature-interdependenz. Technical Report 60, Arbeiten der Fachrichtung Psychologie der Universität des Saarlandes, Saarbrücken, 1980.
- [Zim93] K. Zimmermann. Enhancing qualitative spatial reasoning - combining orientation and distance. In A. U. Frank and I. Campari, editors, *Spatial Information Theory: A Theoretical Basis for GIS*, pages 69–76. Springer, Berlin, Heidelberg, 1993.