ABDUCTIVE REASONING

Abductive diagnosis

Abduction in natural language processing

Weighted abduction

Abduction in natural language generation

ABDUCTIVE DIAGNOSIS

Information

- **Implication (H) – effects of errors**
- **Atoms (S) – symptoms observed**
- **Negated literals (N) – symptoms not observed**

Explanation for S, i.e., minimal errors F, such that F ∪ H ∪ N consistent and F ∪ H |= S

Example

H: measels → fever & rash
migrane → headache & nausea
influenza → headache & sorelimbs & fever

N: empty S: fever → 2 explanations: measles, influenza
fever, rash → 1 explanation: measles
headache → 2 explanations: migrane, influenza
headache, fever → 1 explanation: influenza

N: ¬nausea S: headache → 1 explanation: influenza
headache, rash → 1 explanation: measles & influenza

INTERPRETATION AS ABDUCTION

Sentence interpretation involves

- **Prove the logical form of a sentence together with the constraints that predicates impose on their arguments, allowing for coercion**
- **Merging redundancies where possible**
- **Making assumptions where necessary**

Example

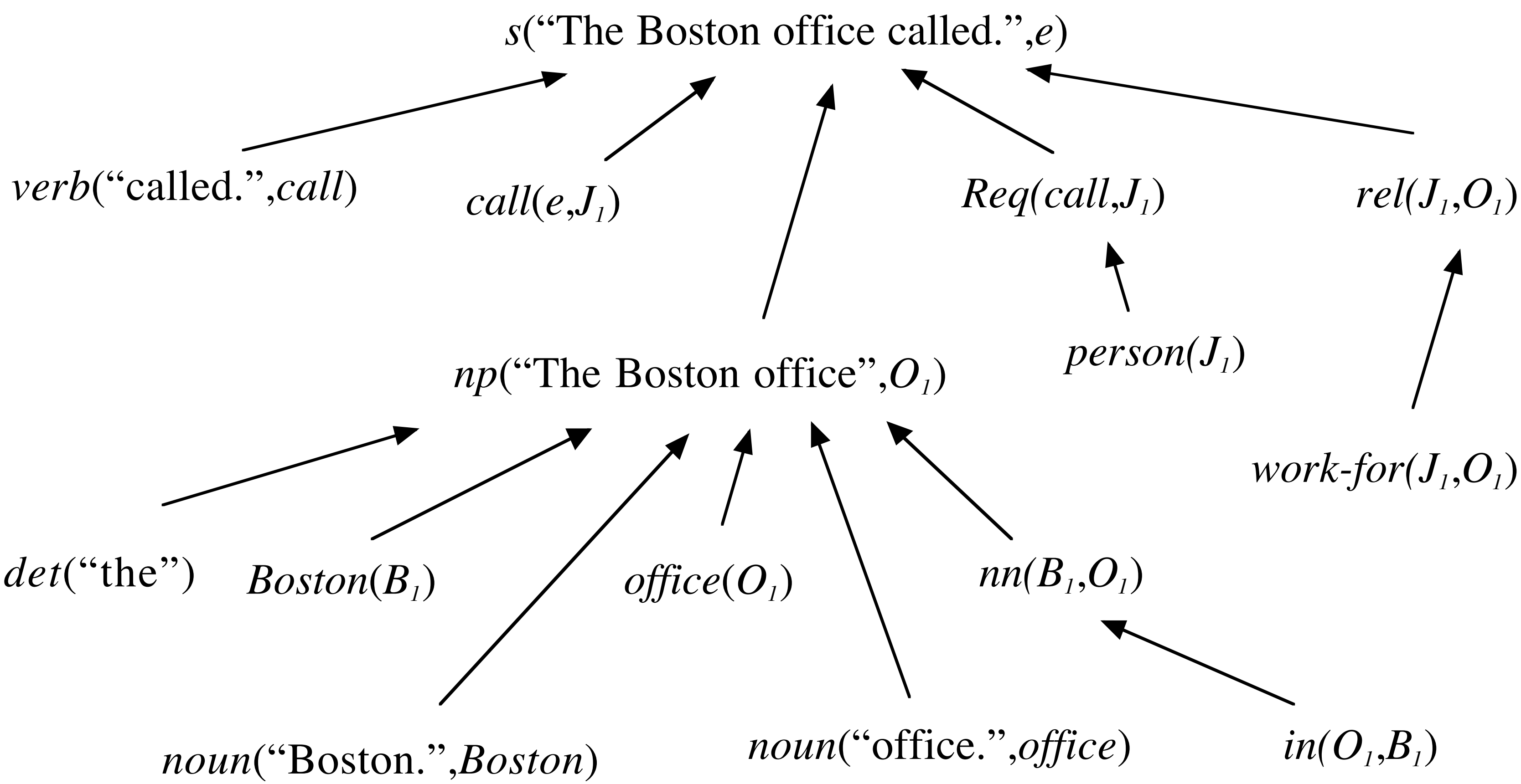
The Boston office called.

- **Resolving reference of "The Boston office"**
- **Expanding the metonymy**
- **Determining the implicit relation between Boston and the office**

Prove abductively the expression

(∃x,y,z,e) call'(e,x) ^ person(x) ^ rel(x,y) ^ office(y) ^ Boston(z) ^ nn(z,y)

INTERPRETATION OF "The Boston office called"



EXAMPLE LOCAL PRAGMATICS

"Disengaged compressor after lube-oil alarm."

To solve local pragmatics, derive the expression

$$(\exists e,x,c,k_1,k_2,y,a,o)\ Past(e) \wedge disengage'(e,x,c)$$
$$\wedge\ compressor(c) \wedge after(k_1,k_2) \wedge event(k_1)$$
$$\wedge\ rel(k_1,y) \wedge y \in \{c,e\} \wedge event(k_2) \wedge rel(k_2,a)$$
$$\wedge\ alarm(a) \wedge nn(o,a) \wedge lube\text{-}oil(o)$$

Interpretation further requires

- **If information for derivation is insufficient, making assumptions is required**
- **Assumptions are new information, with varying likelyhood, according to linguistic form**
- **Costs are assigned to assumptions**

EXAMPLE LOCAL PRAGMATICS (2)

Costs for interpretation

- **Definite NPs generally used referentially**
  Their assumptions should be expensive (10\$)
  The same holds for selectional constraints
- **Indefinite NPs rarely used referentially**
  Their assumptions should be inexpensive (1\$)
- **Indefinite NPs should have intermediate costs (5\$)**
- **Non-nominal propositions are usually new info**
  They should have low costs (3\$)
- **Coercion relations and compound nominal relations require interpretation, hence very high costs (20\$)**

$$(\exists e,x,c,k_1,k_2,y,a,o)\ \mathrm{Past}(e)^{\$3} \wedge \mathrm{disengage}'(e,x,c)^{\$3} \wedge \mathrm{compressor}(c)^{\$5} \wedge$$
$$\mathrm{after}(k_1,k_2)^{\$3} \wedge \mathrm{event}(k_1)^{\$10} \wedge \mathrm{rel}(k_1,y)^{\$20} \wedge y \in \{c,e\} \wedge \mathrm{event}(k_2)^{\$10} \wedge$$
$$\mathrm{rel}(k_2,a)^{\$20} \wedge \mathrm{alarm}(a)^{\$3} \wedge \mathrm{nn}(o,a)^{\$20} \wedge \text{lube-oil}(o)^{\$5}$$

WEIGHTED ABDUCTION

Axioms with costs

$$P_1^{w1} \wedge P_2^{w2} \supset Q$$

If Q is associated with costs c, the costs of assuming P_1 is cw1, and the costs of assuming P_2 is cw2

If w1 + w2 < 1, most-specific abduction is favored

If w1 + w2 > 1, least-specific abduction is favored

Factorization:

$(\exists \ldots x,y,\ldots) \ldots \wedge q(x) \wedge \ldots \wedge q(y) \wedge \ldots$

becomes $(\exists \ldots x,\ldots) \ldots \wedge q(x) \wedge$

if q(x) is associated with lower costs than q(y)

Factorization may override least-specific abduction

$$P_1^{0.6} \wedge P_2^{0.6} \supset Q_1$$

$$P_2^{0.6} \wedge P_3^{0.6} \supset Q_2$$

if we want to derive $Q_1 \wedge Q_2$

ISSUES IN LOCAL PRAGMATICS

Adequate semantic and interpretation requires

- **Reference resolution**

  “The police prohibited the women from demonstrating. They feared violence”
- **Compound nominal interpretation**

  “Lube-oil alarm“, “Boston office”
- **Syntactic ambiguity resolution**

  “Disengaged compressor after lube-oil alarm”
- **Metonymy resolution**

  “The Boston office called”

Further issues include definite reference, distinguishing the given and the new, lexical ambiguity, discourse coherence (rhetorical relations)

APPLICATION AREAS

Source: "Interpretation as Abduction" Hobbs et al., (AI Journal)

TACITUS (The Abductive Commonsense Inference Text Understanding System)

used for message routing, problem monitoring, database entry and diagnosis

Applications so far:

- **Equipment failure reports or casuality reports**
- **Naval operation reports**
- **Newspaper articles and similar texts on terrorist activities**

Example text:

"A cargo train running from Lima to Lorohia was derailed before dawn today after hitting a dynamite charge. Inspector Eugenio Flores died in the explosion. The police reported that the incident took place past midnight in the Carahuaichi-Jaurin area.

TECHNICAL ISSUES FOR FURTHER RESEARCH

Making abduction more efficient

- **Exploiting the domain type hierarchy**
  (for filtering of axioms, axioms for incompatibility of assumptions)
- **Avoiding transitivity axioms**
  (limiting the depth of recursion – e.g., for location containment rules)
- **Reducing the branch factor of the search**
  (for coercion: evaluate class predicates prior to relation predicates)
  (for factoring: avoiding it when this would lead to type violations)

Further pragmatic issues not treated yet:

Resolving quantifier scope ambiguities, metaphor interpretation, recognizing the speaker's plan

A MODERN APPROACH IN THE TACITUS STYLE

Axioms extracted from WordNet and FrameNet, evaluated on textual entailment tasks

John(x1):20 & compose(e1,x1,x2):20 & sonata(x2):20 to be interpreted

Suppose our knowledge base contains the following axioms:

1) form(e0,x1,x2):90 → compose(e0,x1,x2)

2) create art(e0,x1,x2):50 & art piece(x2):40 → compose(e0,x1,x2)

3) art piece(x1):90 → sonata(x1)

I1: John(x1):20 & compose(e1,x1,x2):0 & sonata(x2):0 & cost 56
form(e1,x1,x2):18 & art piece(x2):18

I21 : John(x1):20 & compose(e1,x1,x2):0 & sonata(x2):0 & cost 56
create art(e1,x1,x2):10 & art piece(x2):8 & art piece(x2):18

I22 : John(x1):20 & compose(e1,x1,x2):0 & sonata(x2):0 & cost 38
create art(e1,x1,x2):10 & art piece(x2):8

The “create art” meaning of compose has been brought forward because of the implicit redundancy in the sentence which facilitated the disambiguation

SYSTEM DESIGN FEATURES

Originally a clean and clear reasoning procedure, introducing optimization measures:

- **limitations in time and depth of search**
- **limiting axioms (eg., axioms must "fit" to the input, enable merging with others)**

Exploiting WordNet and FrameNet definitions as axioms

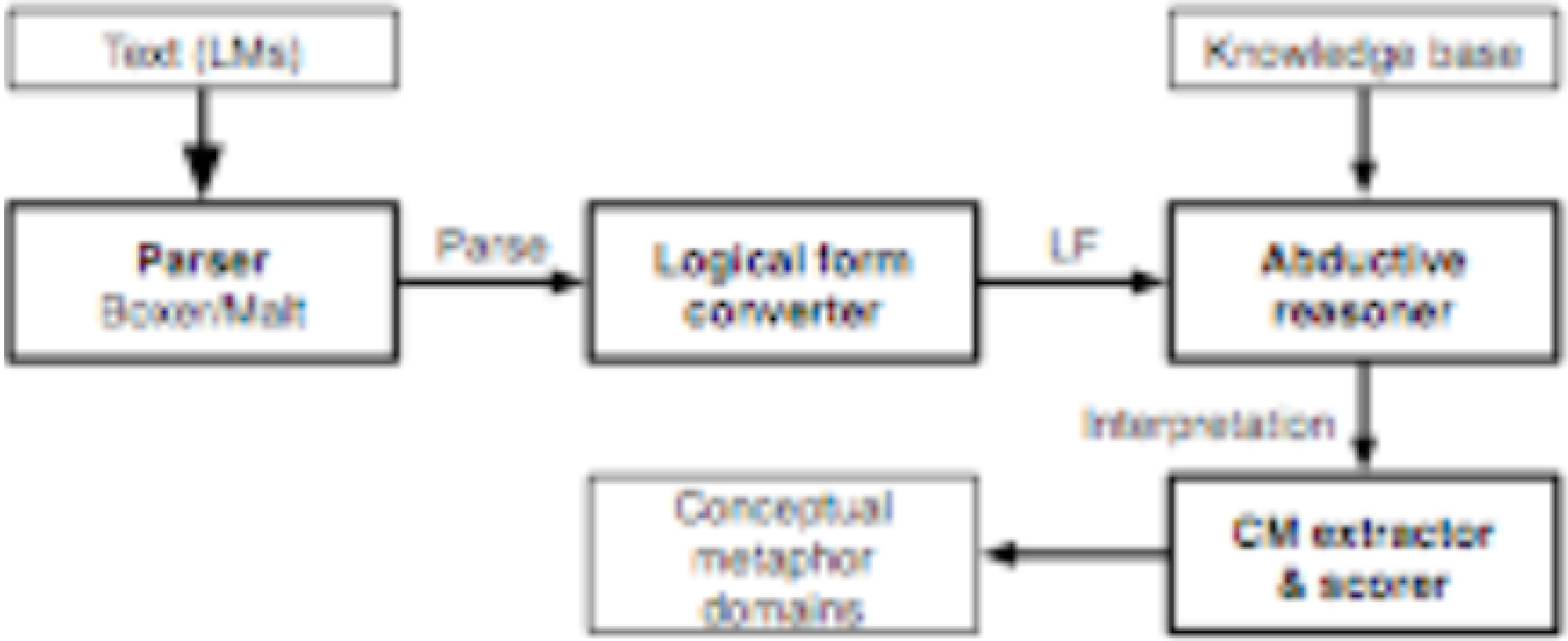
Axiom type	Source	Numb. of axioms
Lexeme-synset mappings	WN 3.0	422,000
Lexeme-synset mappings	WN 2.0	406,000
Synset relations	WN 3.0	141,000
Derivational relations	WN 3.0 (annotated)	35,000
Synset definitions	WN 2.0 (parsed, annotated)	120,500
Lexeme-frame mappings	FN 1.5	50,000
Frame relations	FN 1.5 + corpora	6,000

Interpretation and hypothesis matched

Entailment hypothesized with large cost reductions through factorization and axiom uses

APPLICATION TO METAPHOR INTERPRETATATION

ARCHITECTURE



APPLICATION TO METAPHOR INTERPRETATION DATA

	English	Spanish	Farsi	Russian
Development (sentences)	7,963	8,151	7,349	4,851
Test set (sentences)	894	894	881	644
KB Manual (axioms)	1,595	1,024	1,187	1,601
KB learned (axioms)	3,877	2,558	2,481	3,071

The number of metaphoric sentences in the development and test sets for each language

Mamual and learned knowledge base

APPLICATION TO METAPHOR INTERPRETATATION
RESULTS

	Source		Target	
	Precision	Recall	Precision	Recall
Manual, Spans	79%	22%	84%	49%
Learned, Spans	85%	56%	99%	65%
Hybrid, Spans	81%	57%	99%	70%

Impact of various axiom sources on span-selected Metaphor Mapping performance.

Learned and Hybrid approaches generally outperform the Manual approach to axiom collection

EXPRESSING CAUSAL AND TEMPORAL RELATIONS

The problem

- **Both relations explicit in the knowledge base**
- **Left implicit on the surface, pragmatically inferable**
- **Mimic this to generate natural discourse**
- **Risky, requires adequate model of implicature**

1) Max entered the office. John greeted him with a smile. He showed Max to the seat in front of his desk and offered him a cup of coffee.

2) Max entered the office. Then, John greeted him with a smile. After that, He showed Max to the seat in front of his desk. He then offered Max a cup of coffee.

3) Jon switched off the heating. Judy came in and said the room was too hot.

4) Jon switched off the heating. Then, Judy came in and said the room was too hot.

1) preferred to 2), but 4) preferred to 3)

DISCOURSE CONSTRAINTS

Connection between eventualities (states or actions)

Relations C: causation, part-whole, temporal relation

- ***Temporal Coherence***

  A text is *temporally coherent* if *H* can infer that at least one of the relations *C* holds between the eventualities described in the sentences

- ***Temporal Reliability***

  A text is *temporally reliable* if one of the relations in C which *H* infers hold does in fact hold between the eventualities described in the sentences

Temporally *incoherent*, when no relations inferable
Temporally *unreliable*, when inferable relation false
Temporally *adequate*, when *reliable* and *reliable*

Implicatures calculated via default rules

Manifestations of Gricean-style pragmatic maxims and world knowledge

DISCOURSE RELATIONS (1)

$e_1 \prec e_2$ eventuality e_1 preceeds e_2

$\phi > \psi$ if ϕ then normally ψ

- ***Narration*** **$\{\alpha,\beta\} >$ Narration(α,β)**
  If clauses α and β are discourse-related, then normally *Narration*(α,β) holds

- ***Axiom on Narration*** ***Narration*$(\alpha,\beta) \rightarrow e_\alpha \prec e_\beta$**
  If *Narration*(α,β) holds, and α and β describe events e_1 and e_2 then e_1 occurs before e_1

- ***Explanation*** **$\{\alpha,\beta\}$ ^ cause$(e_\beta,e_\alpha) >$ *Explanation*(α,β)**
  If clauses α and β are discourse-related, and the event described in β caused that described in α, then normally *Explanation*(α,β) holds

DISCOURSE RELATIONS (2)

- ***Axiom on Explanation*** ***Explanation***$(\alpha,\beta) \rightarrow \neg e_\alpha \prec e_\beta$
  If *Explanation*(α,β) holds, then event e_1 described in α does not occur before event e_2 described in β

- ***Push Causal Law*** $\{\alpha,\beta\}$ **^ "fall&push" > cause(**e_β,e_α**)**

  If clauses α and β are discourse-related, and α describes the event e_1 of x falling and β the event e_2 of y pushing x, then normally e_2 causes e_1

- ***Causes precede Effects cause***$(e_2,e_1) \rightarrow \neg e_1 \prec e_2$
  If event e_2 causes e_1, then e_1 doesn't occur before e_2

INTERPRETATION (1)

Assuming that text is coherent

Attempt to connect clauses via discourse relations

$\phi > \psi$ if ϕ then normally ψ

- **Max stood up. John greeted him.**
  ***Narration* inferred**
- **John greeted Max. Max stood up.**
  ***Narration* inferred too**
- **Max fell. John pushed him.**
  ***Narration* and *push causal law* cannot hold both**

  Inferring Narration is *defeasible* here

INTERPRETATION (2)

***Causality modeling*: Keyness as relation on eventualities**

$key(e_1,e_2)$ means e_1 is a key event relative to e_2

$$Narration(\alpha,\beta) \rightarrow (\exists e)(key(e,e_\alpha) \wedge (key(e,e_\beta) \wedge \neg(key(e_\alpha,e_\beta) \wedge \neg(key(e_\beta,e_\alpha)$$

$$Explanation(\alpha,\beta) \rightarrow key(e_\alpha,e_\beta)$$

- **?Max's car broke down. Mary died her hair black.**
  No discourse relation inferable
- **?Everyone laughed. Fred told a joke.**
  Strength of keyness unclear

GENERATION

Knowledge sources

Δ Background knowledge and text purpose

EC Causal and part/whole relations (eventualities)

ET Temporal relations between eventualities

EK Keyness of eventualities

Sketch of the algorithm (three abductive steps)

- **From *EC*, *ET*, *Δ* abduce *EK* fitting speaker's purpose**
- **From *EC*, *ET*, *EK, Δ* abduce discourse structure *D* (depth-first left-to-right on *EC*)**
- **Build *Conc,* concrete assumptions (depth-first left-to-right on *D*)**

a) Add concrete assumptions according to current pair of clauses in *D*

b) Nonmonotonic deductive check on *Δ* and *Conc*
If discourse and event relations included, goto a)
If not, abduce on any rule in *Δ* and add further concrete assumptions about current clause-pair. Goto b)

GENERATION EXAMPLE 1

Two eventualities e_1 and e_2 and e_2 causes e_1

Falling is the key event

EC **{{*fall*(m,e_1), *push*(j,m,e_2)}, {cause(e_2,e_1)}}**

ET **{$e_2 \prec e_1$}**

EK ***key*(e_1,e_2)**

D ***Explanation*(α,β), where *fall*(m,e_α) and *push*(j,m,e_β) D is obtained by abduction of the rules Explanation and Key Event of Explanation**

Conc **{<γ,α,β>,*fall*(m,e_α), *push*(j,m,e_β)}**

NMDC **cause(e_β,e_α) and *Explanation*(α,β)**

- **Either send *Conc* to the surface grammar:**

Max fell. John pushed him.

- **Or abduce some furtehr concrete assumptions:**

<α,β> ^ because(β) > cause(e_β,e_α)

Max fell because John pushed him.

GENERATION EXAMPLE 2

Two eventualities e_1 and e_2 and e_1 preceeds e_2

No key event

EC {{*fall*(m,e_1), *push*(j,m,e_2)}, {}}

ET {$e_1 \prec e_2$}

EK **Neither e_1 nor e_2 *key***

D *Narration*(α,β), *fall*(m,e_α) **and** *push*(j,m,e_β) D **is obtained by abduction of the rules Narration and Key Event of Narration**

Conc {$<\gamma,\alpha,\beta>$,*fall*(m,e_α), *push*(j,m,e_β)}

NMDC **cause(e_β,e_α) and *Explanation*(α,β). The check indicates that the existing set of concrete assumptions will lead to unreliable text. We must therefore add to *Conc* further assumptions about α and β:**

- ***Narration* will not suffice, since it doesn't add any further assumptions to *Conc*; but if we use this rule about and then, we will add a further concrete Assumption.**

- $<\gamma,\alpha,\beta>$ **^ andthen(β) > Narration(α,β)**

Max fell and then John pushed him.

EXPRESSING CONTENT USING GENERALISATIONS

An Example – Necklace in an exhibition of 20th century jewellery

1) **This necklace is in the arts-and-crafts style. It is made of silver, amethysts and pearls. It has very elaborate festoons. It has faceted stones.**

2) **This necklace is in the arts-and-crafts style. It is made of silver, amethysts and pearls. Arts-and-crafts jewels tend to be intricately worked; for instance, this piece has very elaborate festoons. However, unusually for arts-and-crafts jewellery, this piece has faceted stones. Most arts-and-crafts jewels (see for example the jewels in case 8) have cabochon stones.**

More elaborate text 2) preferable

- **More informative and coherent**
- **Non-obvious relations expressed**
- **References to related material**
- **Defeasibility**
  ***arts-and-crafts-jewel*(x) > *intricately-worked*(x)**

DEFEASIBLE RULES

Vague quantifiers

- ***Most Xs are Y*** **for** $X > Y$ ***Few Xs are Y*** **for** $X > \neg Y$

Coherence relations

- ***Exemplification***

  Arts-and-crafts jewels tend to be intricately worked; *for instance*, this piece has very elaborate festoons.
- ***Amplification***

  This piece has very elaborate festoons. *Indeed*, so do most Arts-and-crafts jewels.
- ***Concession***

  Arts-and-crafts jewels tend to be intricately worked, *but* this piece has clean, geometric lines.

User model

- **Not all arts-and-crafts jewellery is elaborate; *for instance*, this piece has quite plain. *Indeed*, many arts-and-crafts jewels are plain.**
- **This piece was *also* designed by Jessie King, *but* around 1910.**
- **If Jessie King designed it, it was probably designed in 1905.**

EXPRESSING DEFEASIBLE RULES

Matching defeasible rules with knowledge base

- **LHS *and* RHS hold for an object**

  Combination is expected, express regularity
- ***Only* LHS holds for an object**

  Exception is worth stating
- **LHS holds for an object, but RHS is unclear**

  Expection is expressed by a hedge

Determining content

1. **Finding simple facts about an object to be described**
2. **Search through rule base for these facts**

   General case of a rule must be introduced by a simple fact, otherwise an inexplicable subject change results; the fact is linked to the generalization via the coherence relation *DEFINITION*. The generalization is linked back to another fact by the coherence relation *EXEMPLIFICATION* (for facts in accordance with the rule) or *CONCESSION* (for facts not in accordance with the rule).

   Multiple facts may be joined to a generalisation
2. **Search mal-rules in user model for misconceptions**

EXPRESSING QUANTIFIERS

NL sentence	*representation*	*description*
All Xs are Y	$X \rightarrow Y$	**indefeasible rule asserted**
No Xs are Y	$X \rightarrow \neg Y$	**indefeasible rule asserted**
Some Xs are Y	$\neg(X \rightarrow \neg Y)$	**indefeasible rule denied**
Not all Xs are Y	$\neg(X \rightarrow Y)$	**indefeasible rule denied**
Most Xs are Y	$X > Y$	**defeasible rule asserted**
Few Xs are Y	$X > \neg Y$	**defeasible rule asserted**
Many Xs are Y	$\neg(X > \neg Y)$	**defeasible rule denied**
Many Xs are not Y	$\neg(X > Y)$	**defeasible rule denied**

Choosing between quantifiers

- **Several may be applicable**
- **For topic introduction, scalar implicature not important (*some*, in case *many* holds)**
- **For handling misconceptions, some forms preferred**
  "Some art-deco jewels have cabochon stones. This jewel is a case in point". not focused
  "Not all art-deco jewels have faceted stones. This jewel has cabochon stones." better