

DFKI Approach to Dialogue Management

Norbert Reithinger, Elsa Pecourt, Markus Löckelt
{bert,loeckelt,pecourt}@dfki.de



What we develop at DFKI



- Research goal:
 - Develop multimodal interactivity interfaces
- Hypotheses:
 - Multimodal interaction facilitates interaction
 - Analysis, dialog management and presentation are tightly connected
- Major task in our group:
 - Design and implement (parts of) the dialog backbone of such systems
- Design approach (our mantras)
 - Use general solutions for general problems
 - No processing and presentation without representation

Dialogue Management

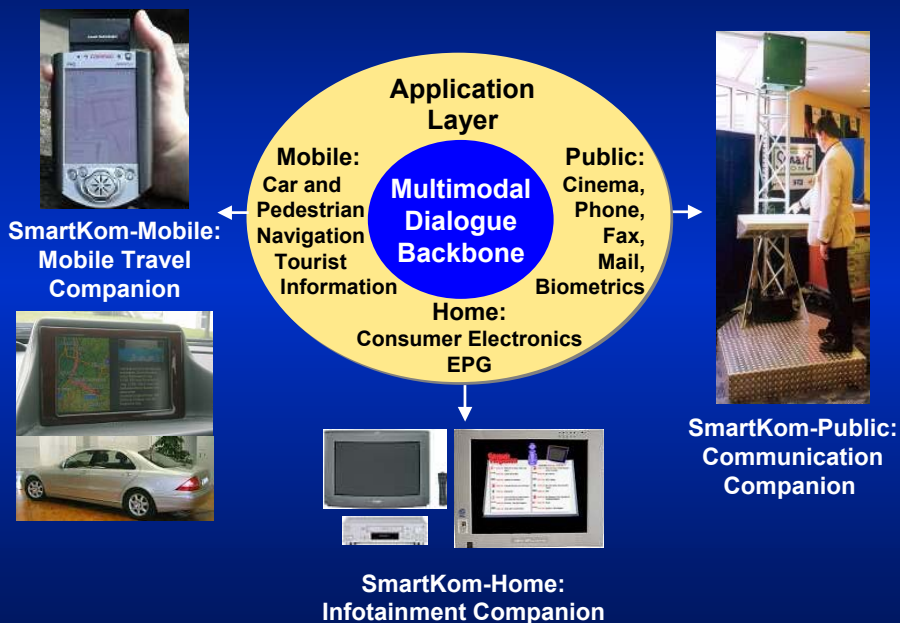
Complex interplay of

- **Action Planner** – the mind of the machine:
 - knows all other acting parties and how to handle them
 - Interaction Types:
 - Request/Response, e.g., DB query, user interaction
 - Active processes, e.g., route guidance
- **Discourse Model(ler)** – the memory of the machine
- **Analysis/Presentation modules**
- **Functionalities**

Current emphasis:

- Flexible and adaptable to many functionalities
- Switch between applications during dialogs
- Reusable (comes with a price)

System1: SmartKom



SmartKom

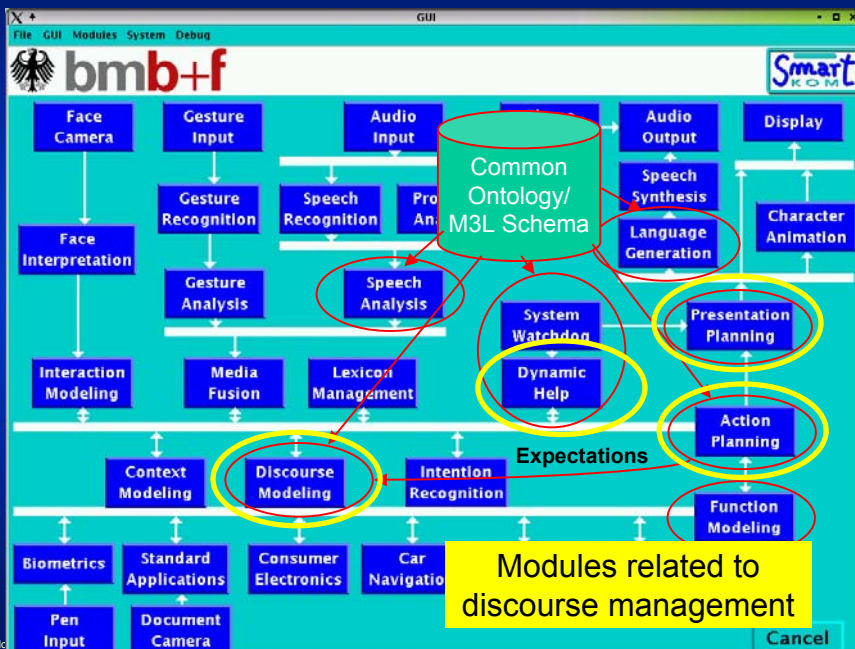
3 scenarios
14 applications
52 functionalities
3500 words at start
3 PCs (2xLinux, 1x Windows)

(... and 12 partner institutions collaborating)

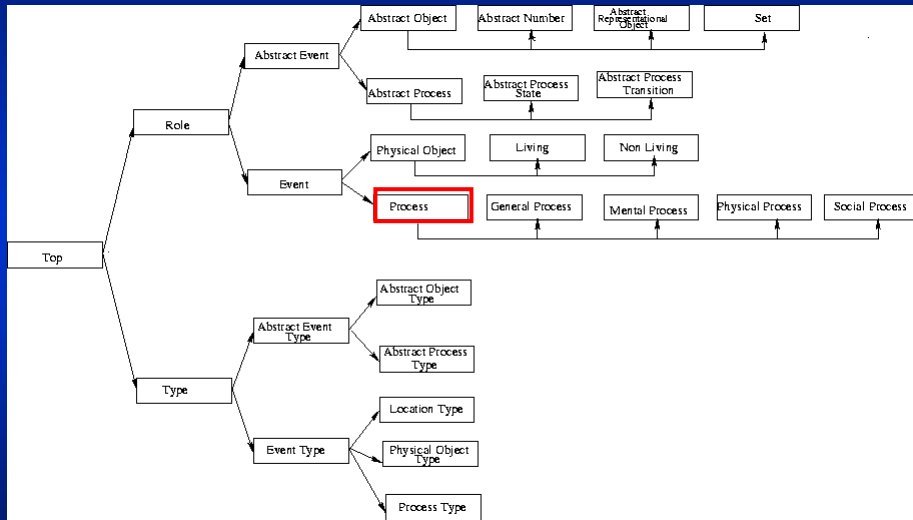
(Source: Reithinger et al: SmartKom - Adaptive and Flexible Multimodal Access to Multiple Applications. In ICM '03)

Home			
EPG (Electronic-Programming Guide)	General program Channel selection Channel information Selection based on genre	Information for one broadcast Time-based operations Help functions for genres	
TV	On/off	Channel selection	7
VCR control	On/off Record Play Pause	Wind/rewind Programming using EPG and the calendar	2
Lean-Forward/ Lean-Backward	Select Lean-Backward Deactivate Lean-Backward	Context aware presentations	6
Total Home			18
Public			
Telephone	Manipulative key operations Telephony functions	Audio handling Address book	4
Hand contour biometry	Selection of biometry type Hand biometry	Presentation and camera control Address book (see above)	3
Voice biometry	Presentation and audio control Voice biometry	Address book (see above) Selection of biometry type (see above)	2
Signature biometry	Presentation and tablet control Signature biometry	Address book (see above) Selection of biometry type (see above)	2
Fax	Presentation and interaction Fax handling	Address book (see above) Camera control	3
E-Mail	Presentation and interaction E-Mail handling	Address book (see above) Camera control (see above)	2
Cinema	General program Movie information	Seat reservation Cinema location	4
Total Public			20
Mobile			
Car navigation	Selection of start und goal city Route type selection Car route computation	Selection of parking garage Information about parking garages	5
Pedestrian navigation	Selection of map type Selection of start und goal Route computation	Selection of points of interest Information for points of interest Integrated car and pedestrian route planning	6
Map manipulation	Resize Help functions for map interactions	Change viewpoint	3
Total Mobile			14
Total System			52

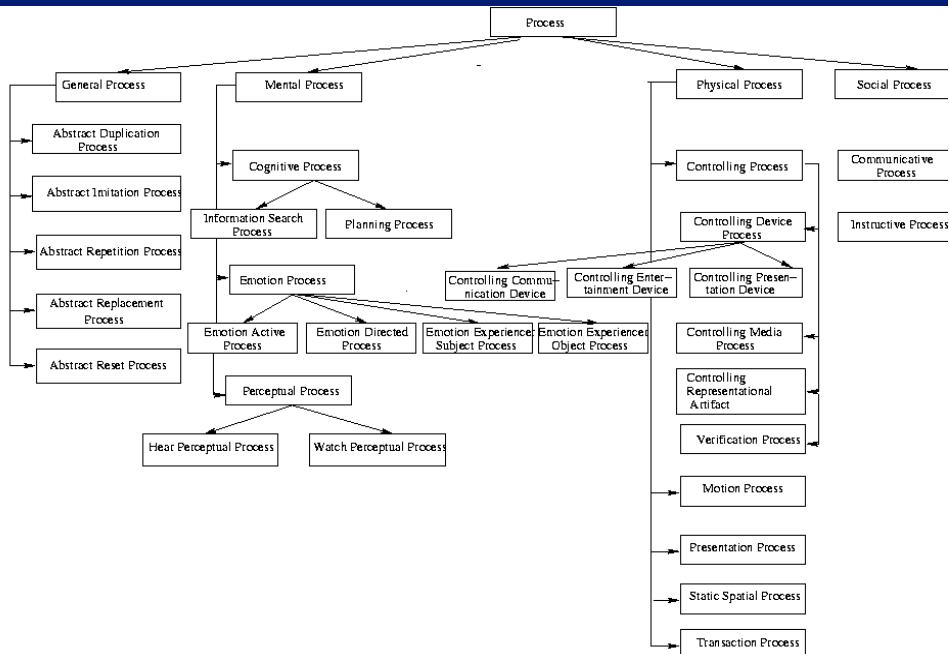
Module Overview of SmartKom



Top Level Ontology

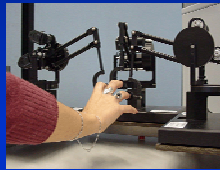


Process Model



System 2: MIAMM – Speech and Haptics

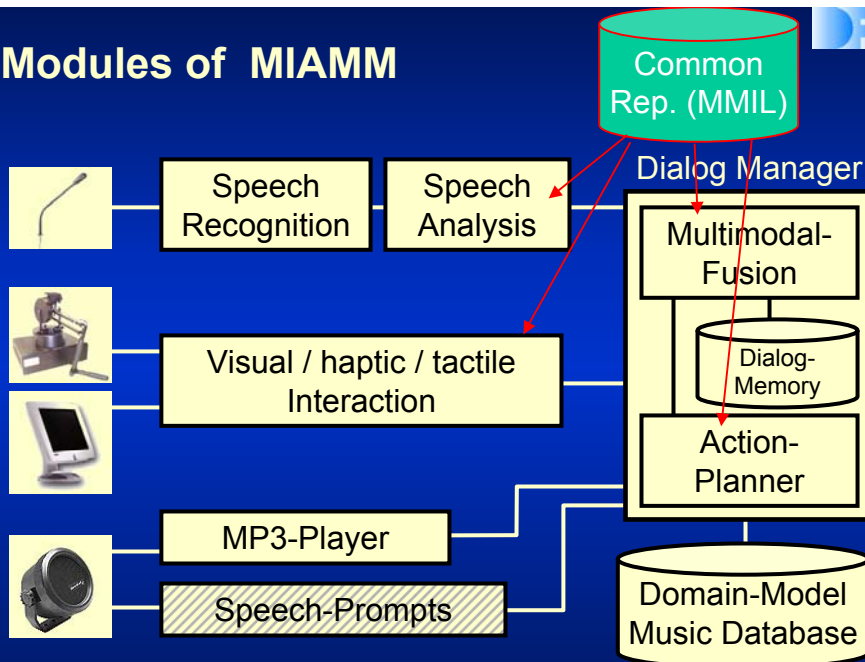
- Application: Search in MP3 Database (DE, FR, EN)
- Haptic buttons enable new interaction
- Simulated with Phantom devices



- Other visualisation metaphors



Modules of MIAMM



Adaptation Needed



- **Representation/Ontology:**
 - Used by almost all modules
 - Provided by partners (SmartKom and MIAMM)
 - Basis of reasoning in the **Discourse Modeller** (SmartKom)
- Interpretation rules in **Speech Interpretation** (SPIN) to create semantic representations
- Plan operators in the **Action Planner**
- **Interface** to the backend application/application logic
 - E.g., we cannot work very well with the prompts/text in the backend
- Functionality that could be skipped:
 - Presentation plans for the **Presentation Planner**
 - Generator rules for the **Text Generator**
- Adaptation time per application in SmartKom: approx. 1-2 person months (2-3 weeks, 2-3 persons)

Approach in the Action Planner: Communication Channels



- Unified treatment of all communication channels (user modalities, internal and external functionalities)
- Input and output channels defined for every module communicating with AP.
 - In MIAMM: MMF, VisHapTac, MiaDoMo, MP3 Player
 - In SmartKom: Intention Recognition, Discourse Modeller, Dynamic Help, Presentation Planning, Function Modelling
- Processing Approach: Communicative Games

Communicative Games (MIAMM)

Interaction

User

Query

request

Presentation

response

„I want something
from the 90's“

„Here are the
results“

Communicative Games

Interaction

User

Extern

Query

Presentation

MiaDoMo (Database)

request

Query

response

Reply

query

results

Communicative Games

Interaction

User

External

Query

Presentation

Clarification dialogue
(System initiative)

Database

request

response

Question

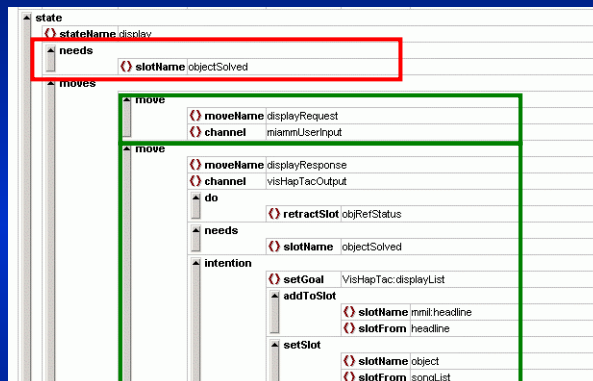
Answer

Query

Reply

„What kind of music
are you looking for“

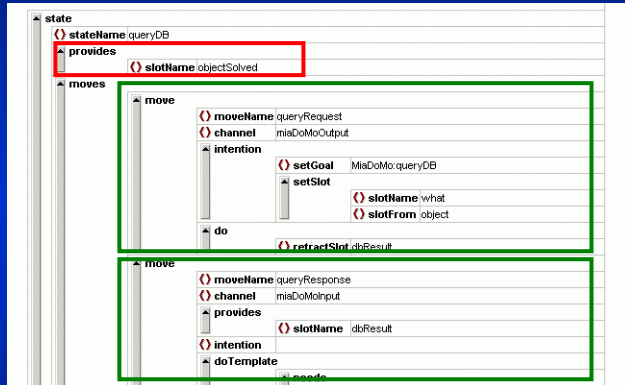
Plan operators in MIAMM



I want something
from the 90's

Here are the
results

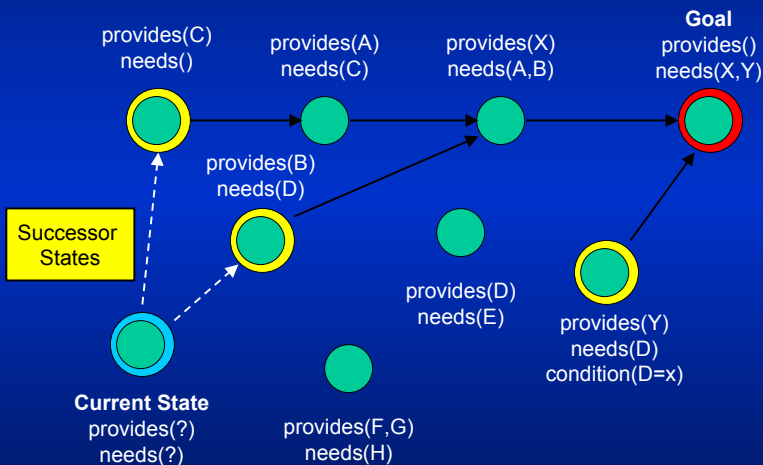
Plan operators in MIAMM (2)



query

results

Planing Approach: Backward Chaining



Pros and Cons of Our Approach

Pros:

- Generic Methods for
 - Analysis/Generation
 - Reference processing
 - Action Planning
- Integration of multimodal I/O
- Neutral to functionalities
- Straightforward method of adaptation

Cons:

- Needs deep(-er) modeling of the domain
- Lots of modules need to go along for a new application
- Difficult to integrate functionalities like the WS-backend which contains e.g. answer texts
- Separation of Dialog History and Action Planner

What Would We Like to Cover

- Error handling in the planner:
 - MIAMM: straightforward solution as plan operators
 - SmartKom: separate Dynamic Help module (due to reasons of project structure)
- Less applications, more generic dialog phenomena
 - More explicit feedback about the systems intentions and the ongoing dialog to the user
 - Dialog processing better adapted to the user
 - Knowledge about argumentations/discussions/negotiations (current bias on access to functionalities)
- Questions:
 - Which abstractions are really needed?
 - Which abstractions are useful?

Thank you very much for your attention!

- Multimodal projects with participation of DFKI
 - SmartKom (BMBF): <http://www.smartkom.org>
 - MIAMM (EU): Multidimensional Information Access using Multiple Modalities: <http://www.miamm.org>
 - COMIC (EU): COntersational Multimodal Interaction with Computers: <http://www.hcrc.ed.ac.uk/comic>
 - VirtualHuman (BMBF): Virtual agents for education <http://www.virtual-human.org>