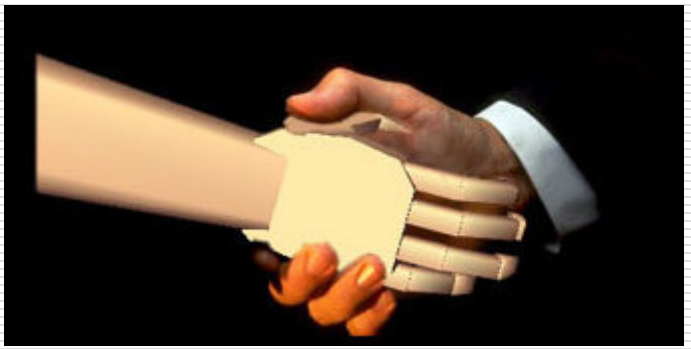
Human-computer interaction

Agent-based interfaces

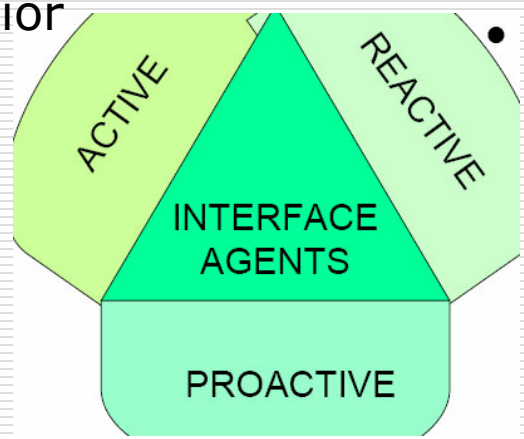


What's an agent?

- They all share ascription of human-like characteristics to software that...
  - presents human-like appearance
  - is autonomous
  - is context-sensitive
  - has specialized expert knowledge
  - learns
  - etc.
- Usually focus on some of these characteristics, the agents don't satisfy all of them

What are agent-based interfaces?

- Appear as life-like characters
- Plan interactive behavior autonomously



- Respond immediately to interruptions
- Handle questions or direct manipulation

- Anticipate the user's needs
- Adopt the user's goals
- Can initiate interaction
- Provide unsolicited comments

Basic terminology

- Agent
  - Autonomous, social, proactive, reactive
  - Employs a certain expertise to support the user in solving a particular problem
- Anthropomorphic agent
  - agent with human-like appearance (cartoon-like ⇨ ... ⇨ realistic)
  - Use of body for communication purpose
- Avatar
  - Bodily representation of a user in virtual worlds
  - Real-time ⇨ user-guided, smart ⇨ autonomous

Agent issues

- Anthropomorphism
  - Do we want to think about the agent as being like a human being?
- Autonomy
  - How much authority do we want to give the agent to act on our behalf?
- Feedback
  - How can we tell what the agent is doing?
- Instructability
  - How can we influence the agent's behavior?

Interface issues

- What can (and should) an agent do?
- How they should do it?
  - Implicit vs. explicit tasking
  - Reporting
  - Activity
- How, when, and why should they interact with the user when doing it?
  - Locus of control
  - Mixed initiative
- Who is responsible when things go wrong?

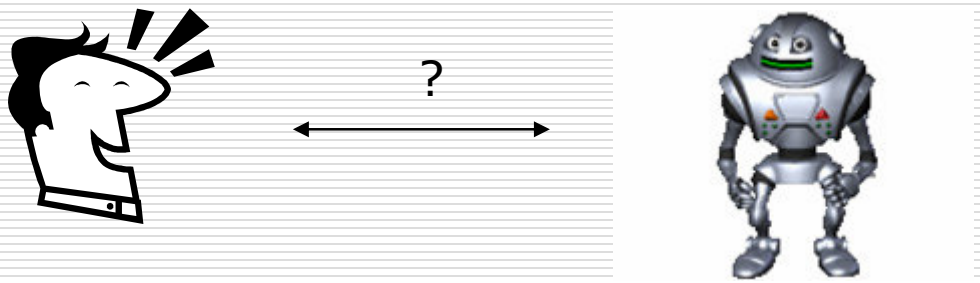
Many kinds of agents

- Software agents for particular tasks
  - Search (e.g., Letiza broadens depth first browsing)
  - Desktop support (e.g. Microsoft's Office Assistant provides Bayesian-based task-sensitive help)
  - Collaborative filters (e.g., shopping recommenders [Resnick and Varian 1997])
- Interface agents, e.g., PPP
- Embodied conversational agents, e.g., REA
- Tutoring agents, e.g. Steve
- Collaborative agents, e.g. Max
- Social, relational agents, e.g. Laura
- Robots, e.g. Leonardo

General value of agents

- Decrease task complexity
  - Bring expertise to the user (in the form of expert critiquing, task completion, co-ordination)
- Provide a more natural (i.e., anthropomorphic) environment with which to interact, e.g.,
  - multimodal communication
  - task, application, and discourse status via facial displays
- Tangible metaphor of mediation
  - There is “somebody” who helps me

how do agents come across? – some results



Effectivity/effeciancy

- Redundancy of multimodality is not always advantegous (Weidenmann, 1997)
- Agents make educational software more effective, simply because children spend more time in front of it (Lester et al., 2000)
- Virtual agents can replace written instructions (Bente)
- Not always facilitation of memory by multimodal output (Krähmer et al.)
- Many open questions
  - systematic analyses are scarce
  - A lot of factors seem to be influential (task, appearance, behavior, user traits, …)

Acceptance

- Graphical interfaces bring about higher acceptance (Hubona & Blanton, 1996; Ahern, 1993)
- Showing an anthropomorphic agent leads to judging the system as more entertaining (Takeuchi & Naito, 1995; Koda & Maes, 1996; van Mulken et al., 1998)
- Perceived intelligence and trustworthiness is increased (King & Oya, 1996; Sproull et al., 1996; Rickenberg & Reeves, 2000)
- Language processing abilities of the system rated higher

Acceptance

- Role of appearance
  - Social evaluation and attribution of friendliness or liking seems to be highly dependent on the specific appearance (Dehn & van Mulken, 2000; Sproull et al., 1996; Koda & Maes, 1996)
  - Controversy about best kind of appearance: Cassell & Thórisson advocate anthropomorphic appearance, Ball & Bates not
  - Parke (1991) recommends a not too realistic appearance since expectations may be raised
  - Agents that resemble user in appearance, gender, ethniticity, etc. rated higher

Acceptance

- Role of (nonverbal) behavior
  - Even subtle nonverbal behaviors influence acceptance of a virtual agent
  - Attribution of sociality and socio-emotional reactions elicited (Bente, Krähmer et al.)
  - Rickenberg & Reeves (2000):
    It is not sufficient „*to focus on whether or not an animated character is present. Rather the ultimate evaluation is similar to those for real people – it depends on what the character does, what it says and how it presents itself*“ (p. 55).

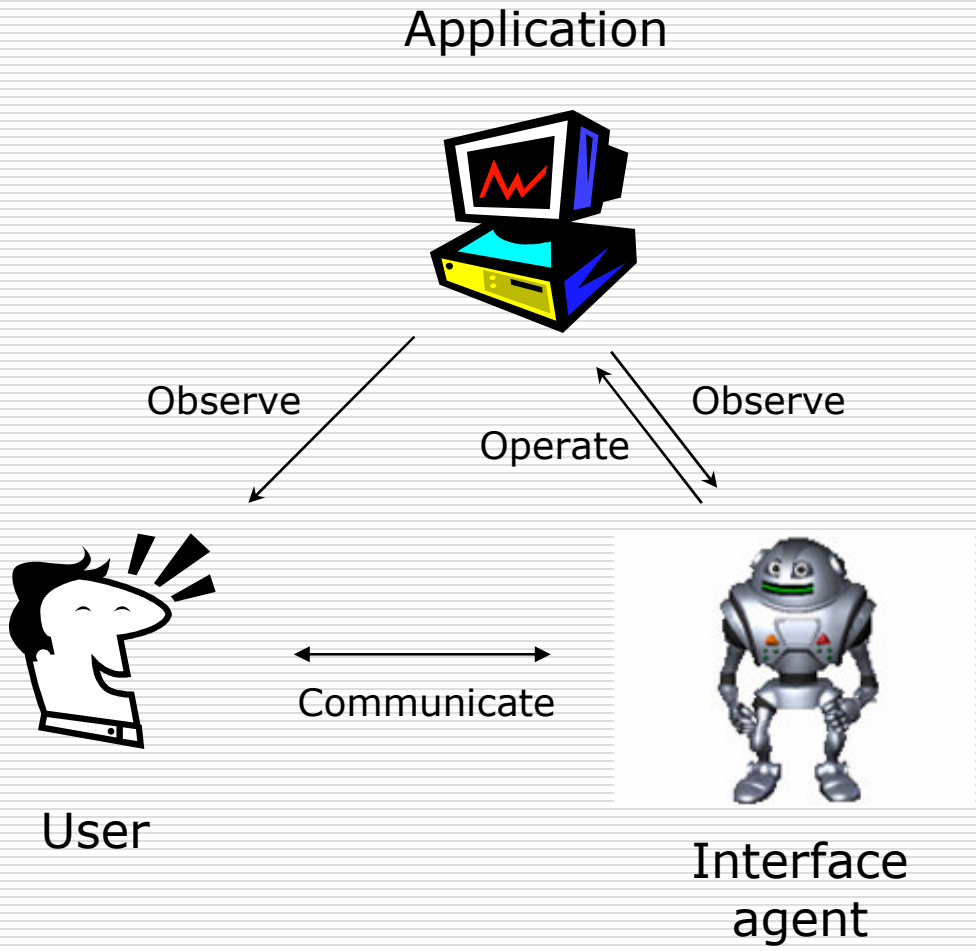
Reactions of the users

- Social presence of the virtual character
  - If confronted with an embodied interface agent, users try to present themselves in a more positive light (Sproull et al., 1996) ⇨ „impression management“
- Effects of social facilitation/social inhibition
  - Task-performance is inhibited by the social presence of a monitoring agent (Rickenberg & Reeves, 2000)
- Attention of the user is drawn to the face
- Open questions
  - Long-term effects, or do humans get used to it?
  - „Many people want computers to be responsive to people. But do we also want people to be responsive to computers?“

Kinds of agent-based interfaces

Interface agents

- Mediate between user and application
- Communicate with the user
- Operate the application for the user
- Agent and user can observe the application



Planned Presentation Persona (PPP)

- Interface agent for the WIP system
- Generates
  - technical instructions
  - Product presentations on the web
- Focus on behavior planning: given a presentation task and some parameters, decide…
  - Which material to present?
  - How to present it?
    - Which acts to perform?
    - Which temporal order?
  - How to achieve a believable overall behavior?

(DFKI, 1994)

Overview of presentation planning

presentation task

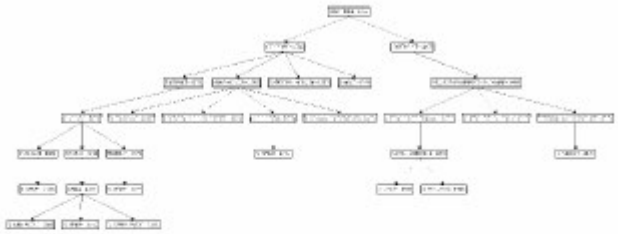
determine presentation acts

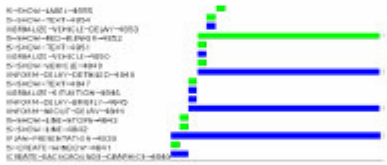
determine presentation schedule

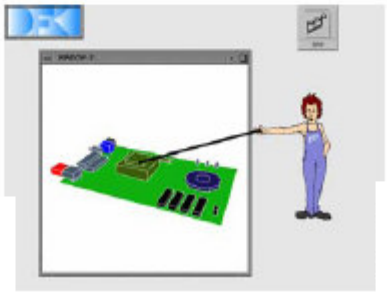
Script 

executed presentation

e.g., describe modem







Embodied agents

- Anthropomorphic appearance
- Different modalities with different benefits: *facial display*, *gaze*, *gesture*, *speech*, *intonation*, *body posture*
  - ➔ Adaptability
- Exchange of information on multiple levels in parallel
  - ➔ Modality synergy
- Natural communication „protocolls“
  - ➔ Increased naturalness, efficiency, smoothness, robustness of communication

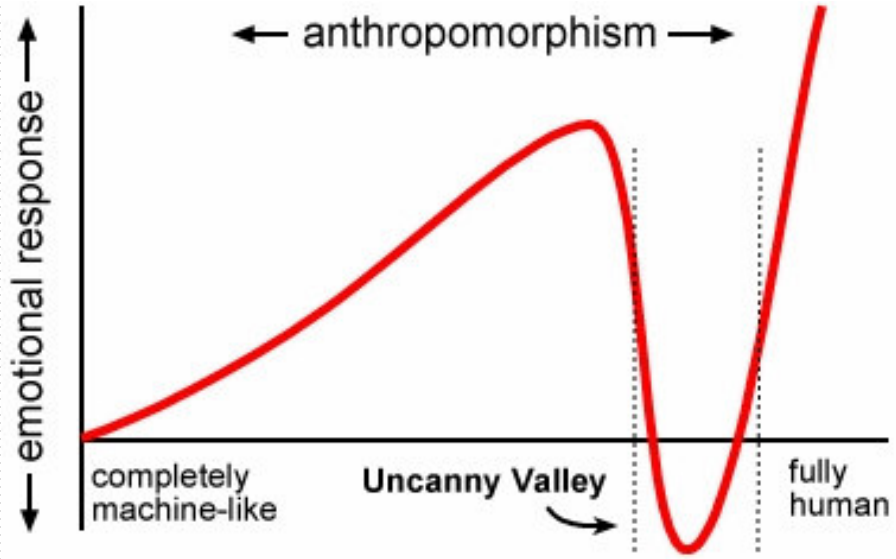
The „Uncanny valley“

Masahiro Mori (late 1970s): Emotional responses to robots vary with anthropomorphism in appearance & motion

- increasingly positive and empathic until suddenly strongly repulsive
- approaches human-human empathy when indistinguishable from humans

Fequent explanation: either human-like or non-human characteristics stand out, generating either empathy or „alienation“

Human-like appearance necessitates human-like behavior!



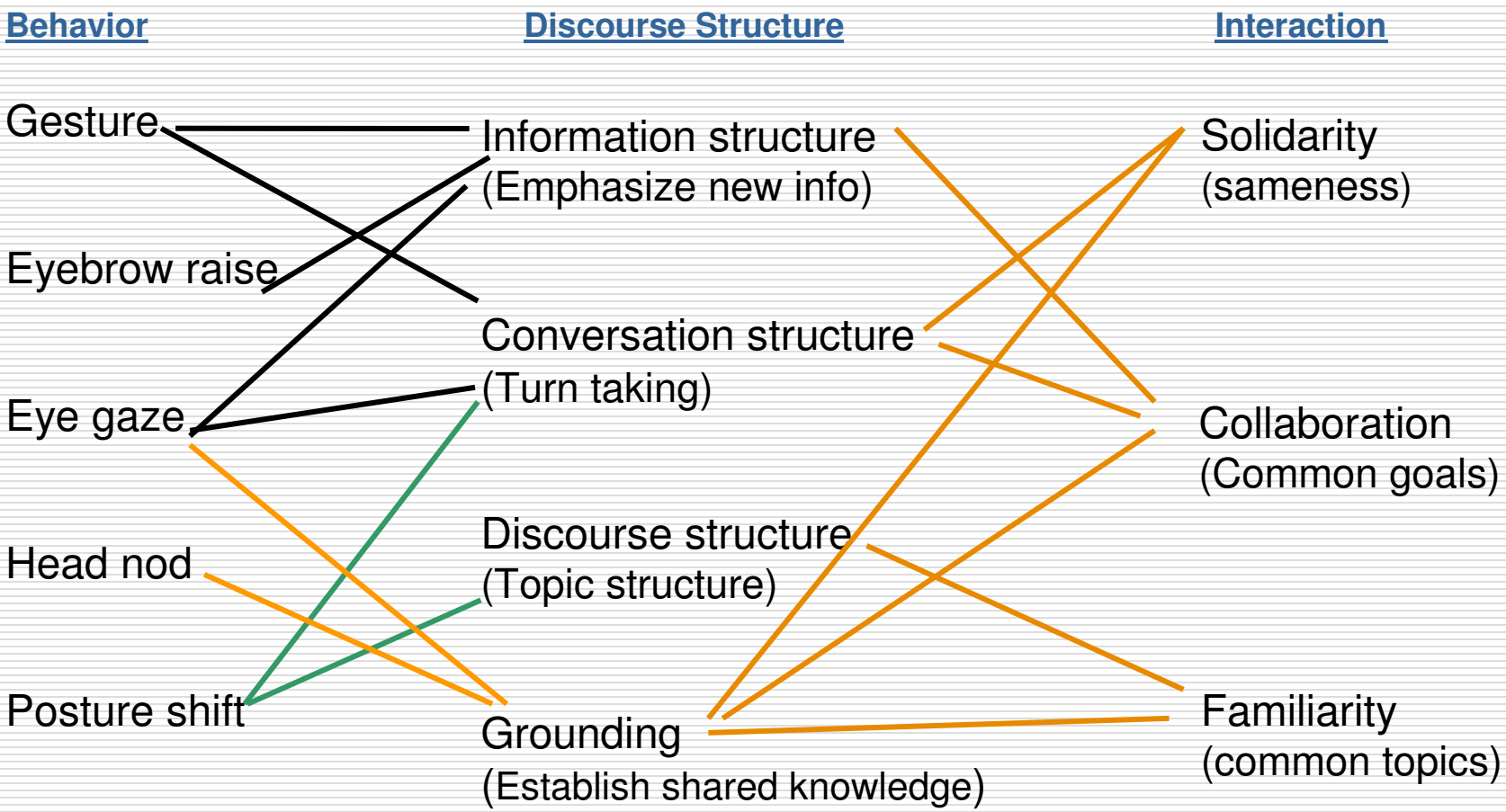
Embodied conversational agents

„Computer interfaces that hold up their end of conversation, have bodies and know how to use it for conversational behaviors as a function of the demands of dialogue and of emotion, personality, and social convention."

(Cassell, 2000)

- Same properties as humans in face-to-face communication
  - Recognize and respond to verbal and nonverbal input
  - Generate verbal and nonverbal input
  - Deal with conversational functions of behaviors (e.g. turn taking, feedback)
  - Participate actively in discourse

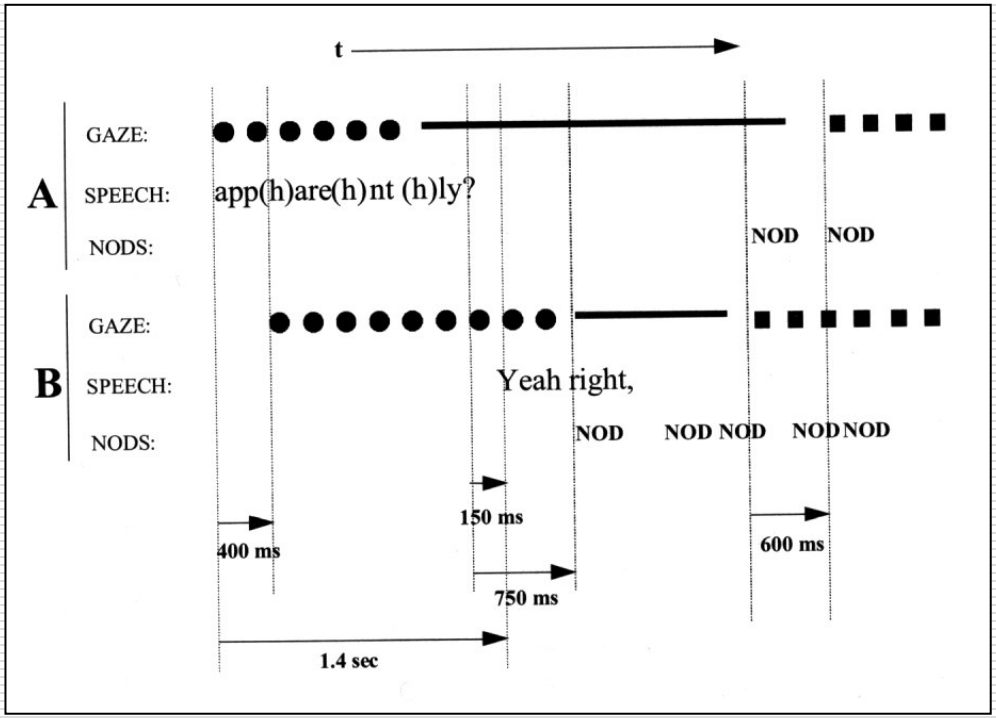
The importance of nonverbal behavior



Functions, **M**odalities, **T**iming, **B**ehavior

- Distinction between
  - propositional and interactional functions of conversation
  - conversational functions and communicative behaviors

- Use of modalities to pursue multiple communicative goals in parallel
- Timing among behaviors on various timescales



Models of conversational function

Interactional and propositional *goals*

- ⇨ conveyed by conversational *functions* (cf), e.g., inivitation, turn taking, turn keeping, provide feedback, emphasize
- ⇨ carried out by communicative *behaviors* (cb)

- A cb may convey several cf's; a cf may be realized by different sets of cb's
- Example: Turn taking (Cassell et al., 2000)

Conv. function	*Comm. behavior*
Give turn	Look, raise eyebrows
Want turn	Raise hands
Take turn	Glance away, start talking

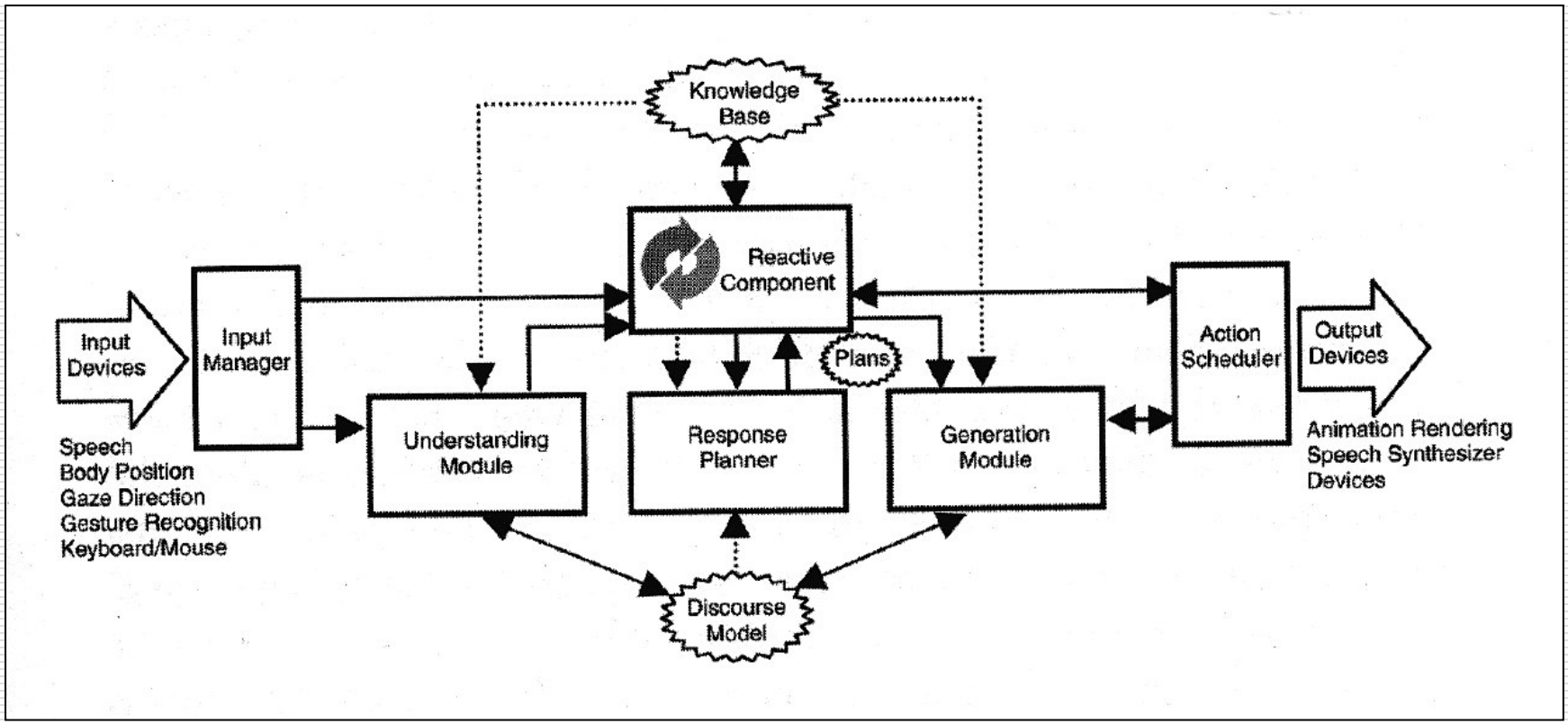
ECA architecture: constraints

- Conversational function model
  - Explicit representation of cf as basis for core operations
  - Repository of cb's
  - Modules for mapping cf on output (cb's) and infering cf's from input (⇨ symmetric architecture)
- Handle propositional and interactional information
  - User model
  - Planning of multi-sentence output
  - Domain and environment knowledge (static, dynamic)
  - Dynamic discourse model

ECA Architecture: constraints

- Multistep deliberation, parallelism, modularity
  - Modules for input understanding, response/dialogue planning, and output generation
  - Processessing in parallel
- Timing & efficiency
  - Handle different threads of communication with different response-time requirements (e.g. quick reaction time for interactional behaviors)
  - Different processes concentrating on activities at various timescales
- Output synchrony

FXPAL architecture



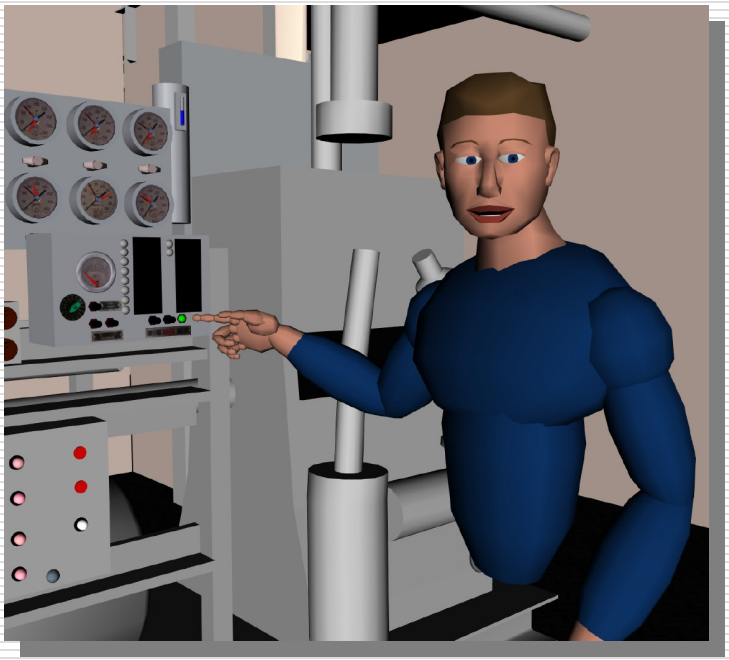
REA – the real estate agent

- Embodied Conversational (interface) agent
- Real estate agent
- FMTB model on top of FXPAL architecture
- Multimodal input/output & active dialogue management

(Cassell et al., 1999, 2000)

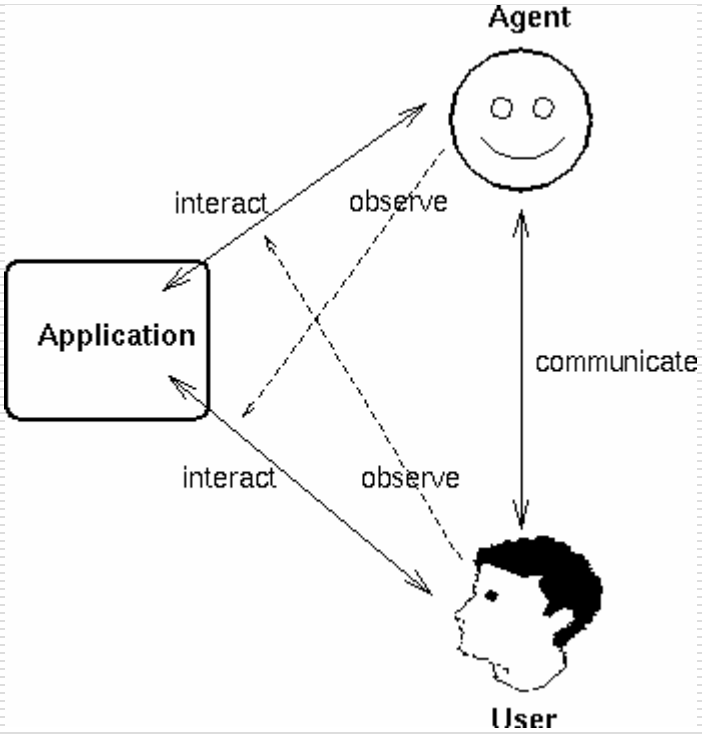
Tutoring (pedagogical) agents

- Teach user interactively a task
- Example: Steve (Rickle & Johnson, 1998, 2000)
  - Teach students how to operate and maintain gas turbine engines aboard naval ships
  - Co-situated with the user in a 3D virtual world
  - Task-oriented dialogues
  - Locomotion & multimodal behaviors (face, gaze, head movements, arm and hand movements)



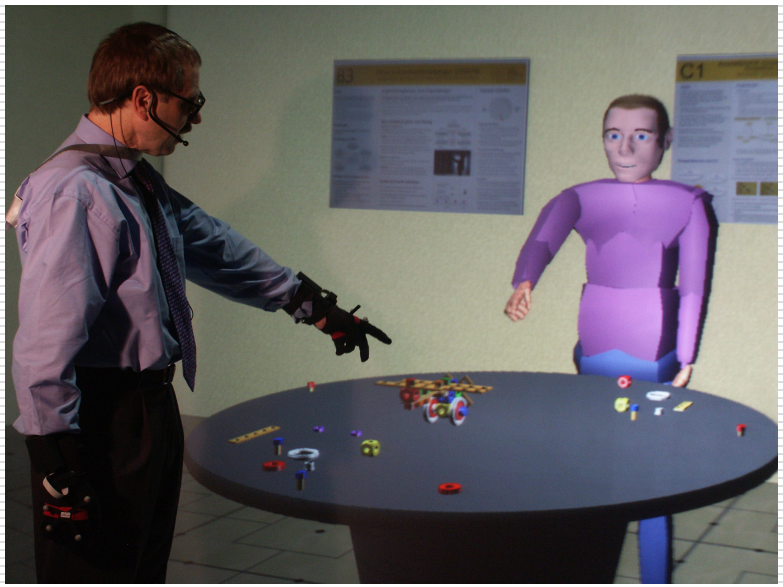
Collaborative agents

- Human and agent collaborate on a task
- Both observe the other
- Both can communicate about the task



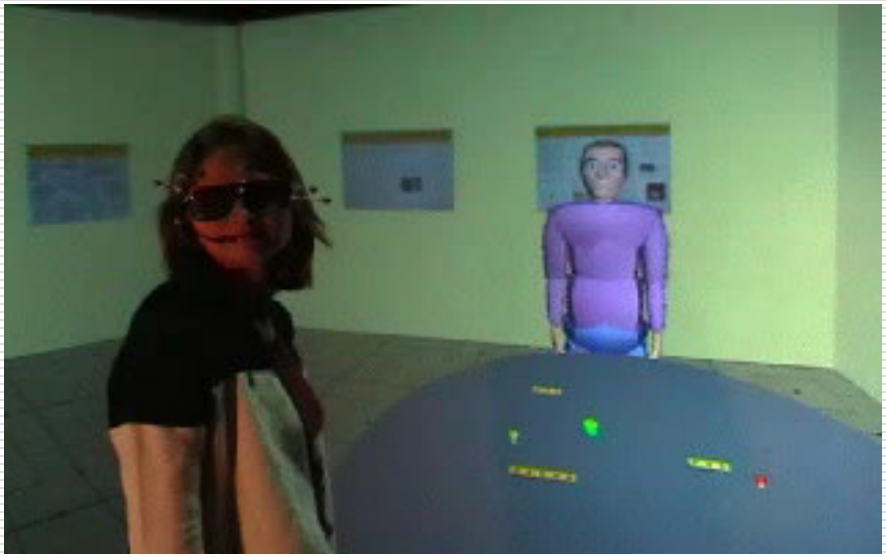
Max (AG-WBS, Uni Bielefeld)

- Collaborates with user in a shared virtual world
- Baufic construction task
- Based on cognitive architecture
  - Auditory and (simulated) visual perception
  - BDI-based deliberation
  - Hybrid architecture
  - On-the-fly utterance generation



(Kopp et al., 2001;
Leßmann & Wachsmuth, 2003;
Kopp & Wachsmuth, 2004)

Max – the assemby expert

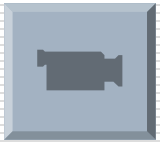


Max – the museum guide

Max – the „receptionist“

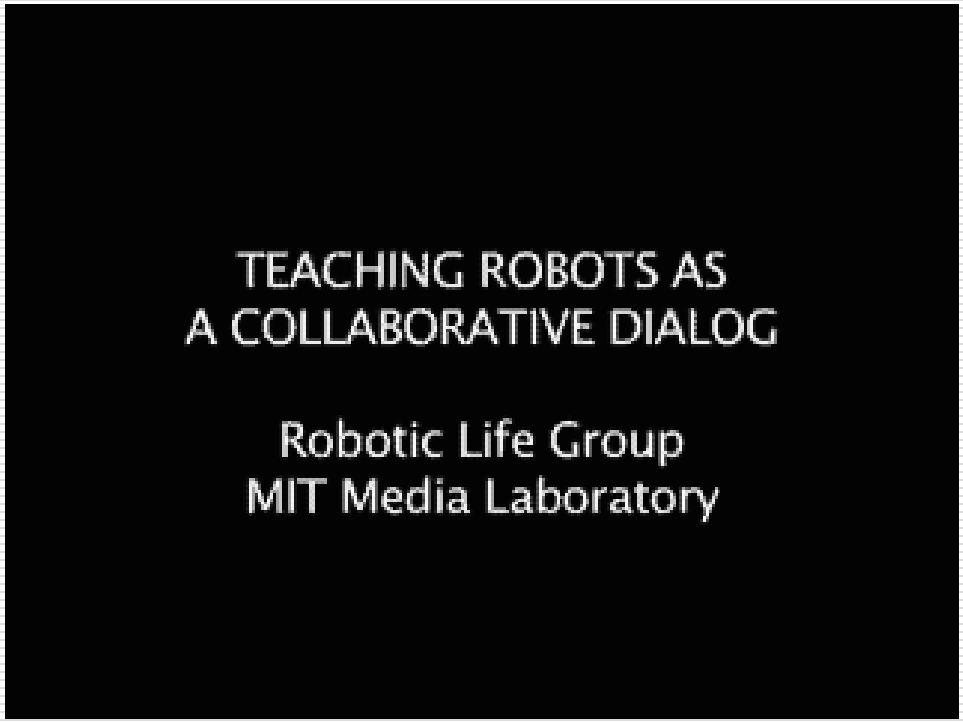
Social agents

- Use human social interaction protocols in the interface
  - Small talk, mirroring to build rapport
  - Immediacy to show liking
  - Flattery to increasing liking
  - etc.
- Relational Agents: Computational artifacts designed to build and maintain long-term, socialemotional relationships with their users
- Example: Laura (Bickmore, Picard, et al.)
  - Small talk, humor, politeness
  - Emotions and empathy
  - Reasoning and communication about relations



Social robot companions: Leonardo

- Collaborative dialogue
- Modeling collaborative behavior by means of social cues
- Able to read same cues from human user and to provide mutual support back (intention recognition)
- Learning goal-directed actions through imitation and vision



C. Breazeal (MIT)

Where are we heading?

- A vision from the Morpha project
  (BMBF; 1999-2003)