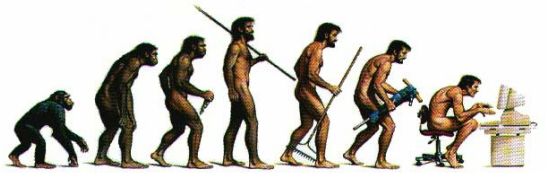
Human-Computer Interaction

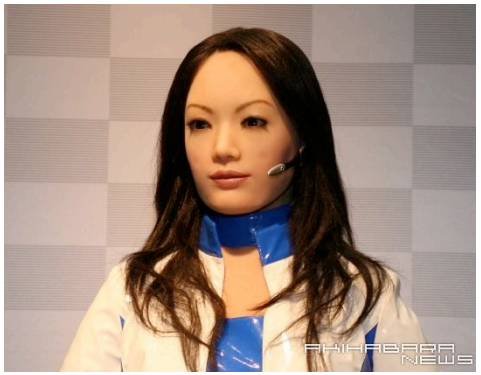
Termin 12:
Agent-based interaction

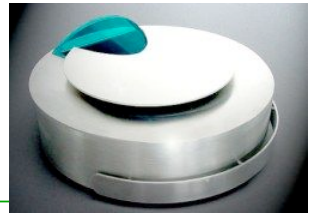
The evolution of user interfaces



Year	Paradigm	Implementation
1950s	None	Switches, punched cards
1970s	Typewriter	Command-line interface
1980s	Desktop	Graphical UI (GUI), direct manipulation
1980s+	Spoken Natural Language	Speech recognition/synthesis, Natural language processing, dialogue systems
1990s+	Natural interaction	Perceptual, multimodal, interactive, conversational, tangible, adaptive
2000s+	Social interaction	Agent-based, anthropomorphic,social, emotional, affective, collaborative

Agent-based interaction





Basic terminology

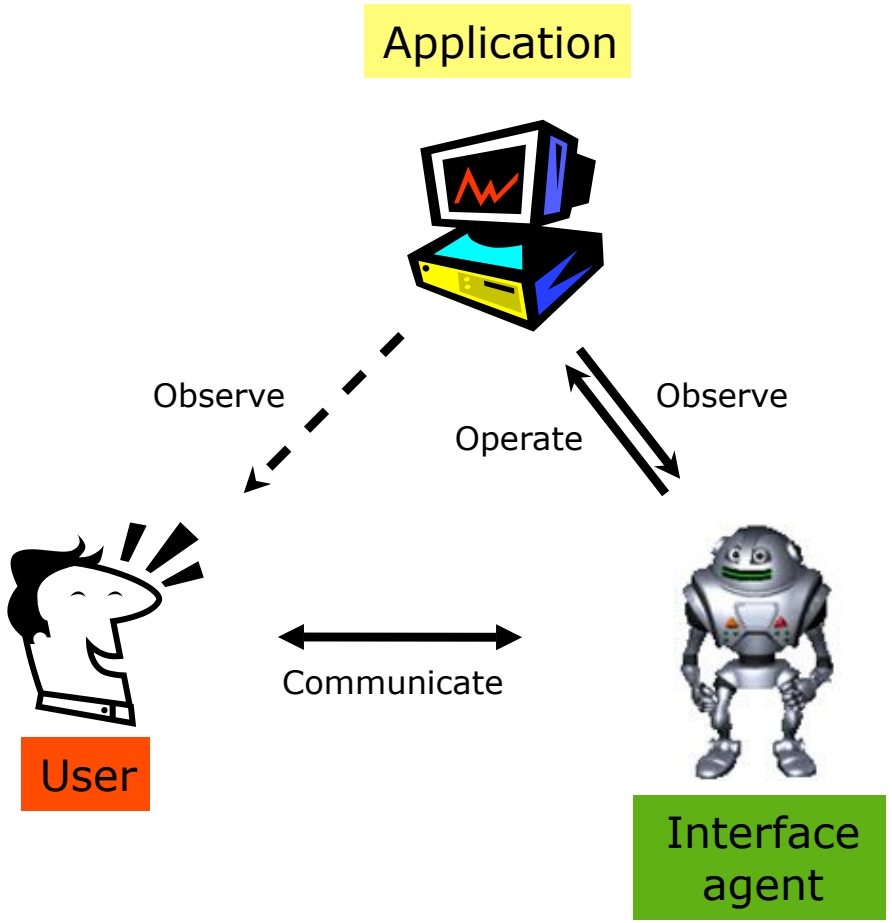
- Agents share ascription of human-like characteristics to software:
  - is social or communicative to other agents
  - is autonomous, reactive, or proactive
  - is context-aware or situated
  - has *specialized* expert knowledge, solve *special* tasks
  - learns, adaptive
- In practice, the term „agent“ is used broadly
  - 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. email), shopping recommenders, auction bots

Basic terminology

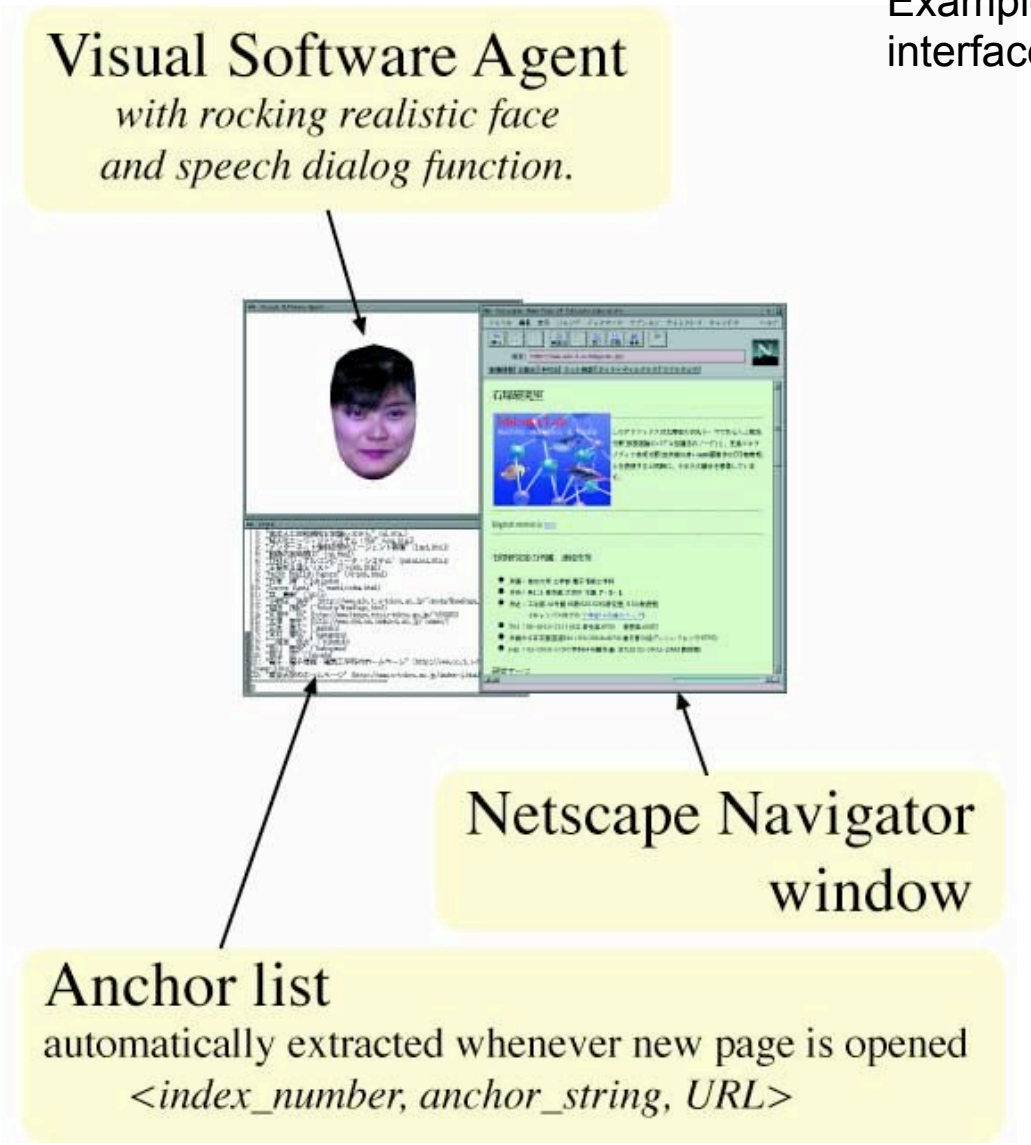
- Avatar
  - Representation of a user in virtual worlds
  - Real-time ⇨ user-guided, smart ⇨ autonomous
- Agent
  - Piece of soft/hardware that is autonomous, social, proactive, reactive
  - Employs a certain expertise to support the user in solving a particular problem
- Anthropomorphic agent
  - agent with human-like appearance (cartoon-like ⇨ ... ⇨ realistic)
  - Use body for communication as well as manipulative purposes

Agents as interfaces

- Task: Mediate between the user and an application
- Communicate with the user
- Operate the application for the user
- Agent and (maybe) the user „observe“ the application



Example: anthropomorphic interface agent (U. Tokyo)

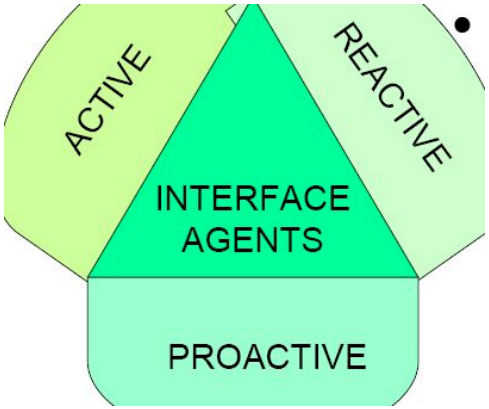


Agent-based interfaces - motives

- 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 in which to interact
  - Should facilitate interaction since people are accustomed to this form of interaction
- Tangible metaphor of mediation
  - There is “somebody” who helps me
- Entertaining & motivating
  - It‘s fun and motivating working together with somebody

Hallmarks of 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

Embodied agents

- Bodily appearance that affords natural *output modalities*: facial display, gaze, gesture, speech, intonation, body posture
- Exploit advantages of natural multimodal communications
  - ➔ Adaptability
  - ➔ Modality synergy
  - ➔ Natural communication „protocolls“
  - ➔ Increased naturalness, efficiency, smoothness, robustness of communication
- Virtual humans

Embodied conversational agents (ECA)

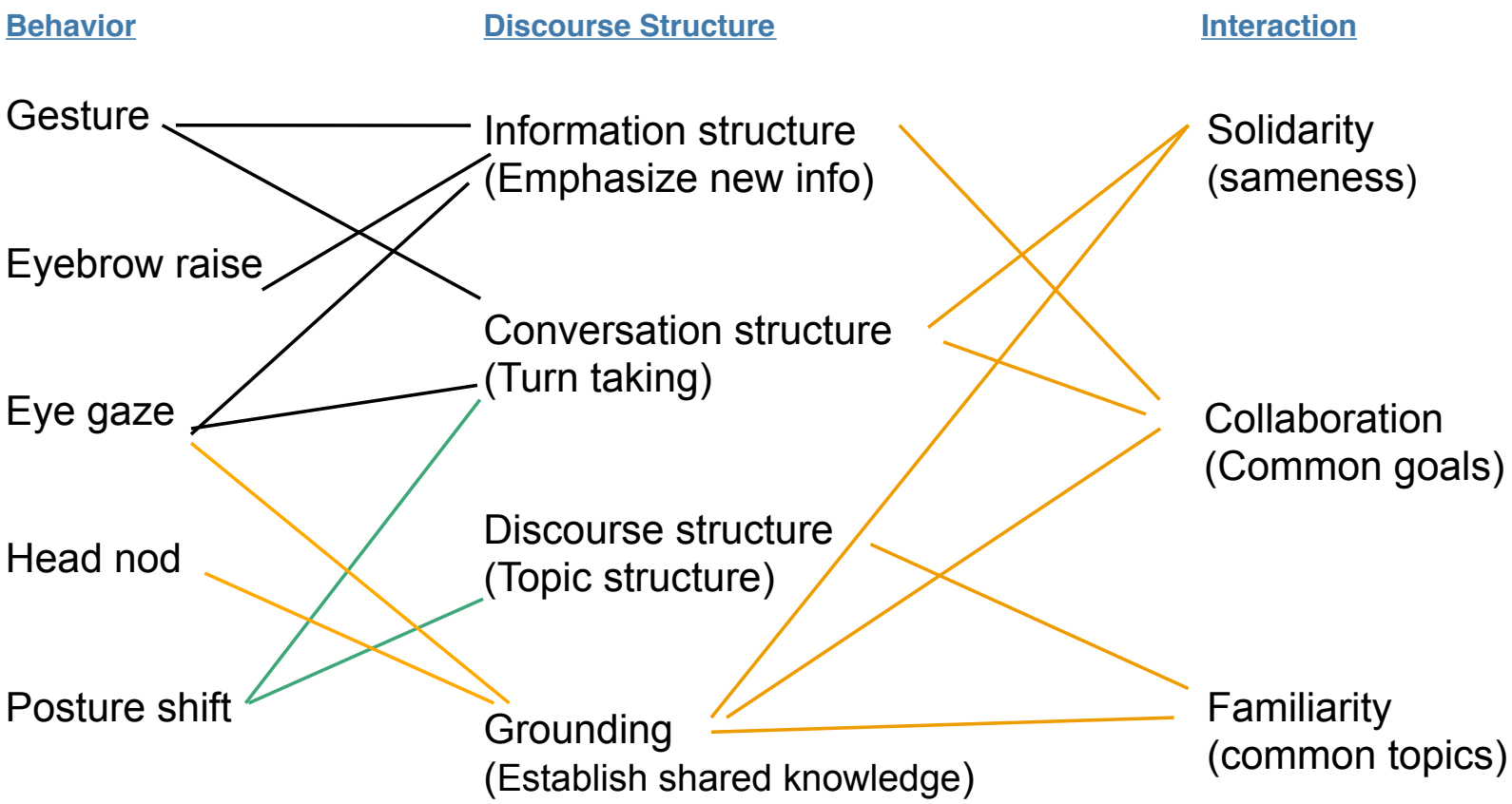
„Computer interfaces that hold up their end of conversation, have bodies and know how to use it for conversational behaviors as a function of the demands of dialogue and of emotion, personality, and social convention.“

(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

Example of conversational agents

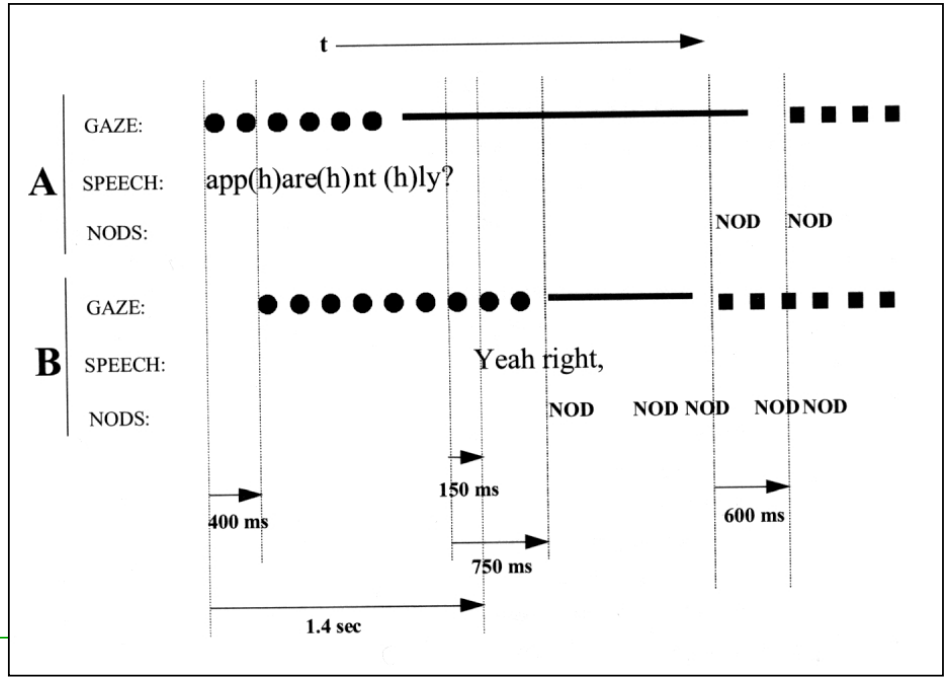
The importance of nonverbal behavior



What's key in multimodal behavior?

- **F**unctions, **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*

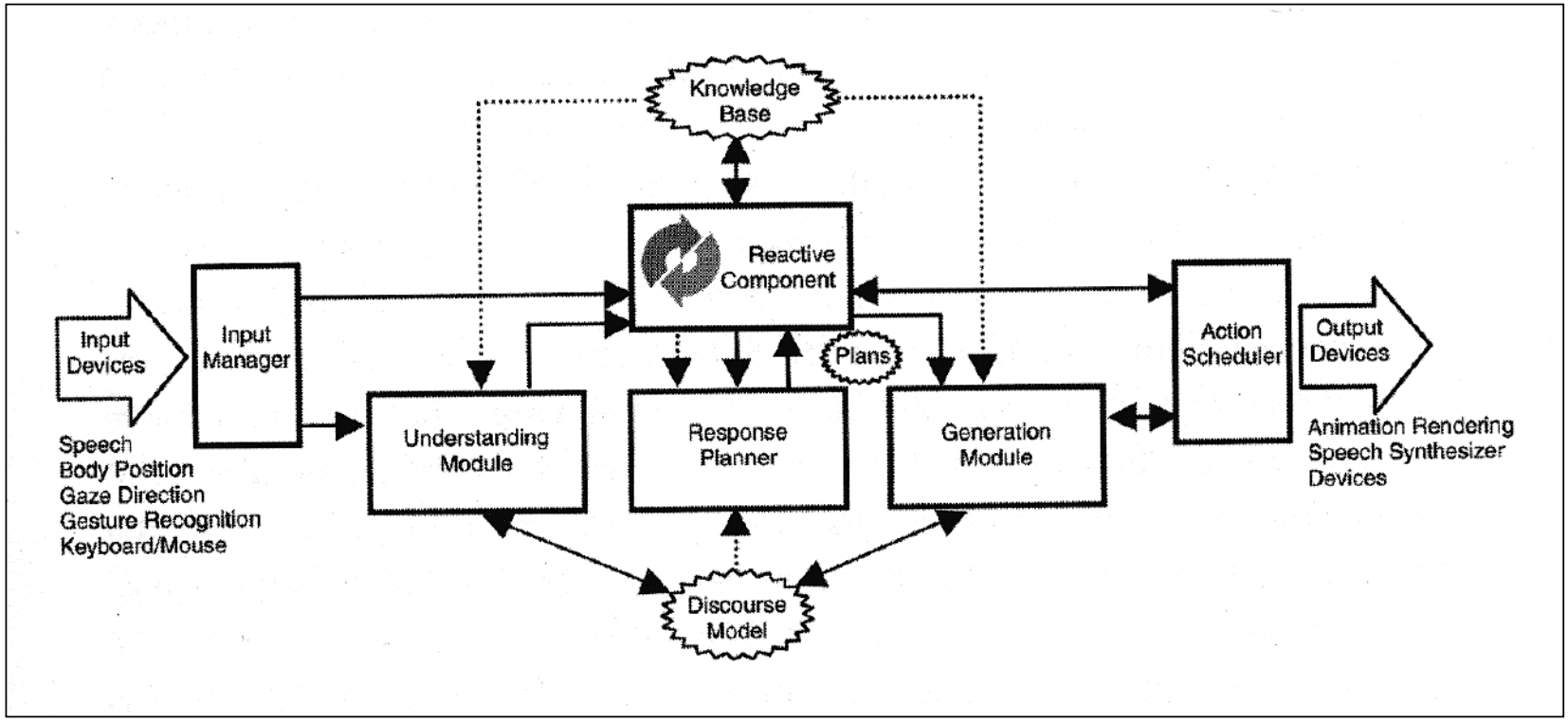
⇨ conveyed by conversational *functions* (cf), e.g., inivitation, turn taking, turn keeping, provide feedback, emphasize

⇨ 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

FXPAL architecture

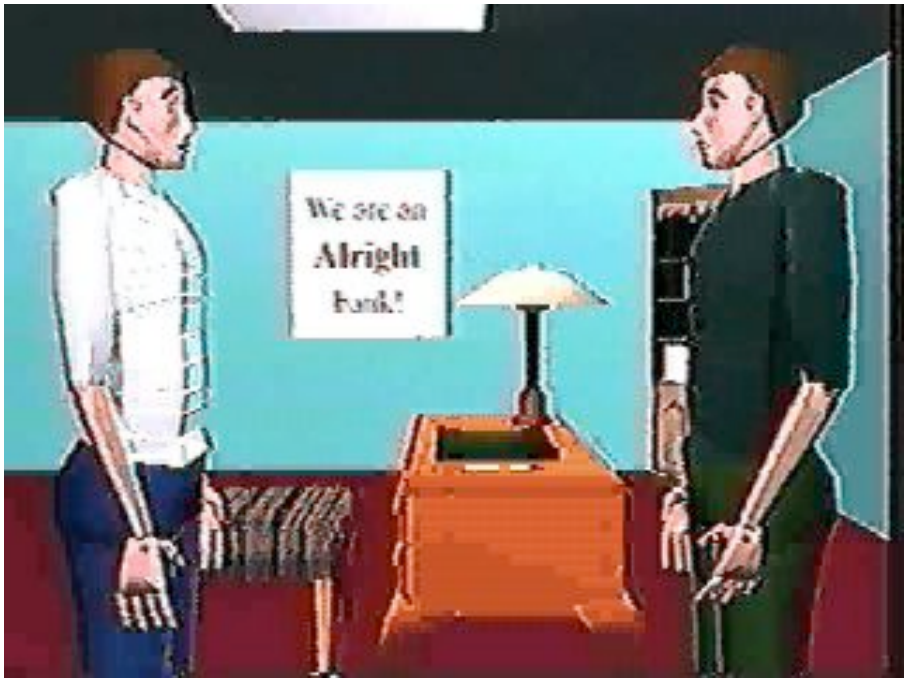


Why is it so hard to build an ECA?

- Conversational function model
  - Explicit representations of cf and cb's and their mappings needed fot both input processing and output generation
- Propositional and interactional information
  - Handling both kinds of information at the same time requires rich dynamic models of user and discourse, as well as large domain and environment knowledge
- Multistep deliberation, parallelism, modularity
  - Input understanding, response/dialogue planning, and output generation must run fast, correct and in parallel
- Timing & efficiency
  - Different threads of communication must be handled at different timescales
- Output synchrony

The beginning of multimodal agents...

- Animated Conversation (1994)



Example: REA (MIT, 2000)

- Scenario: Real estate agent
- Multimodal input/output & active dialogue management

(Cassell et al., 1999, 2000)

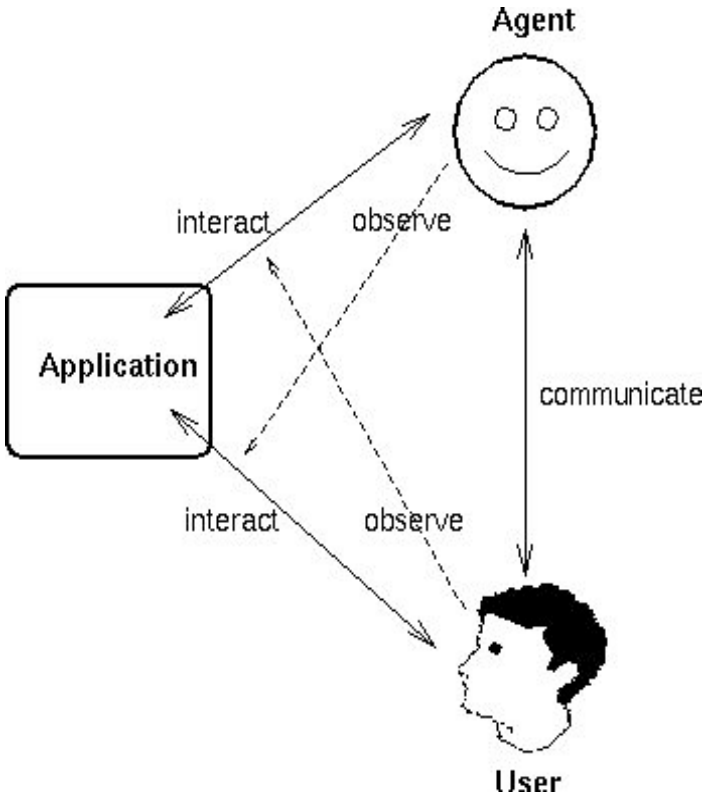
Example: Max (AG-WBS, Uni Bielefeld)

- As conversational museum guide in the HNF

Max at the M4 hallway

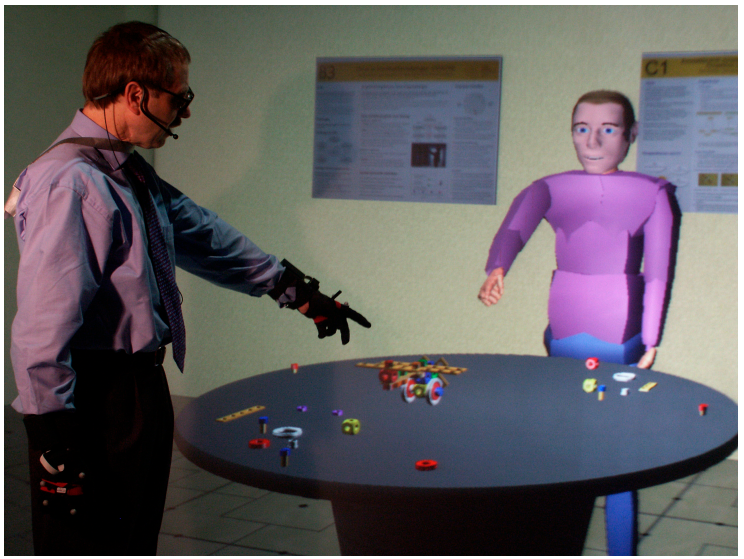
Collaborative agents

- Human and agent *collaborate* on a task
- Both can *actively contribute* to the task
- Both *observe* the other
- Both can *communicate* about the task and their collaboration

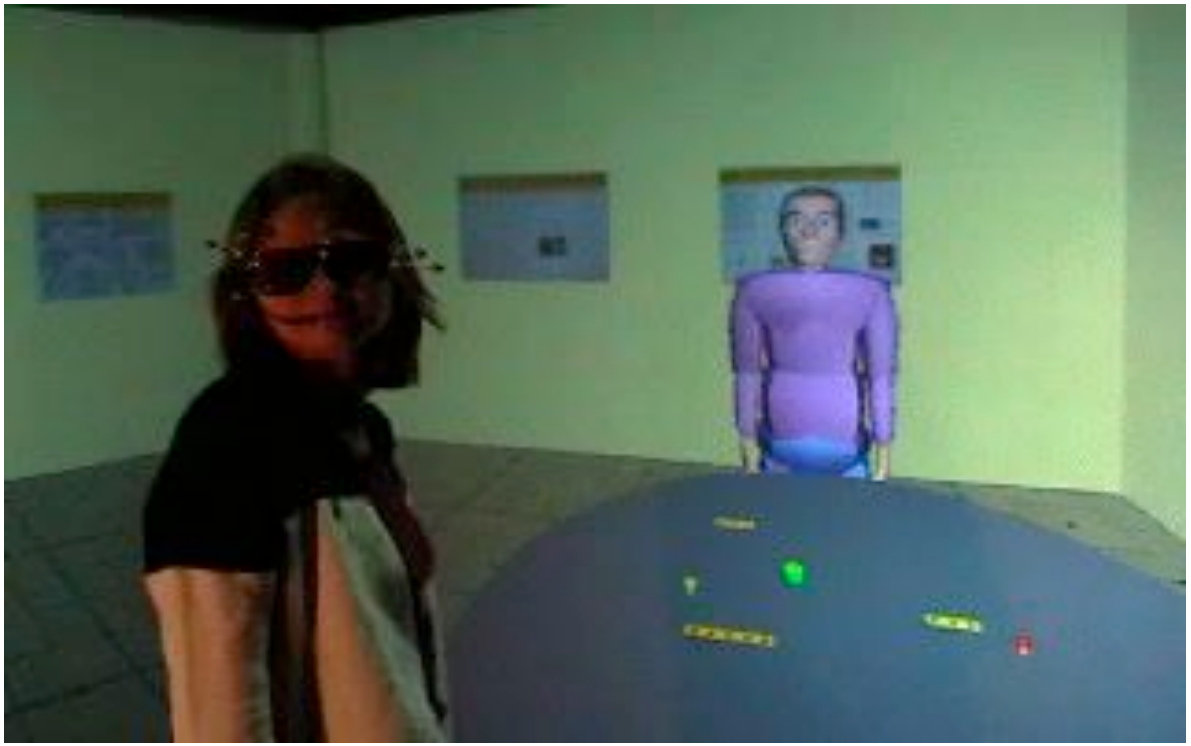


Max (AG-WBS, Uni Bielefeld)

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



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



Agent-based interfaces from the user perspective...

- 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?
- Communication
  - How can we *communicate* with the agent?
- Feedback
  - How can we tell what the agent *is doing*?
- Instructability
  - How can we *influence* the agent's behavior? Locus of control?
- Responsibility
  - Who is liable when things go wrong?

Do anthropomorphic interfaces help?

- Virtual face attracts **attention** (Dehn & van Mulken, 2000).
- Human-like faces cause evaluations as **more entertaining** (Takeuchi & Naito, 1995; Koda & Maes, 1996; van Mulken et al., 1998, Krämer et al., 2002).
- Perceived **intelligence** and **trust** in a system (as well as its credibility) is increased when an anthropomorphic interface is used (Sproull et al. 1996; Walker, Sproull & Subramani, 1994; Rickenberg & Reeves, 2000).
- Users are more inclined to **delegate tasks** to the system when a human like face is visible (Milewski & Lewis, 1997). Other studies could not prove an increased readiness to delegate a task (Krämer & Bente, 2005).
- By means of a robot autistic children are **prompted to interact** with the artifact and with each other (Werry et al., 2001).

Do anthropomorphic interfaces help?

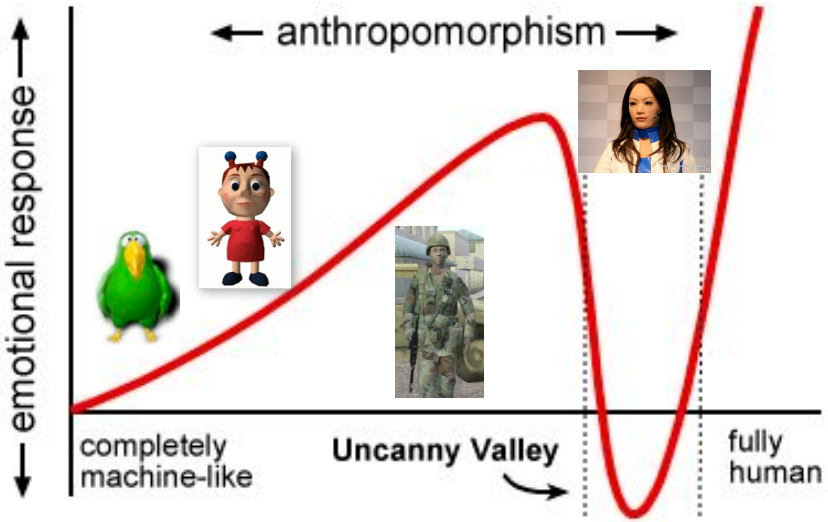
- Graphical interfaces bring about higher **acceptance** (Hubona & Blanton, 1996; Ahern, 1993)
- Role of **appearance**
  - Social evaluation and attribution of friendliness highly dependent on the appearance (Dehn & van Mulken, 2000; Sproull et al., 1996; Koda & Maes, 1996)
  - Controversy: Cassell & 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. are rated higher
- Agent **behavior** must match realism of its appearance

The „uncanny valley“ hypothesis

- how human-like an appearance?

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

Human-like appearance necessitates human-like behavior!

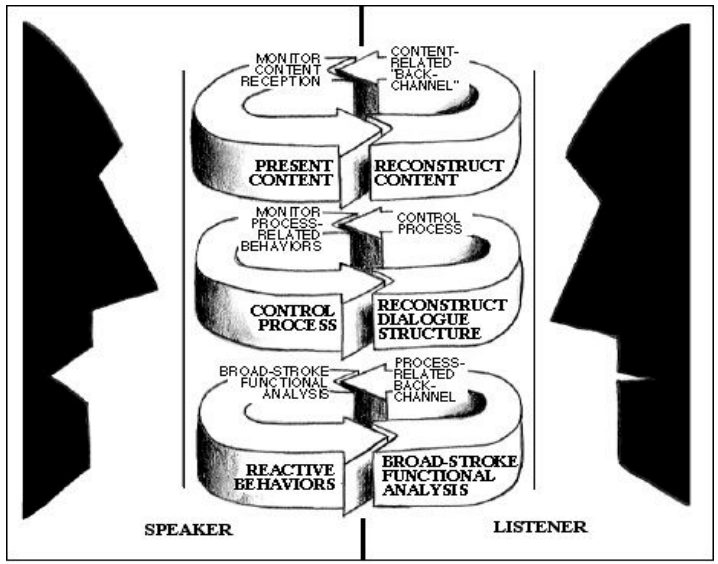


Numerous roboticists questioned Uncanny Valley's scientific status, noting that "we have evidence that it’s true, and evidence that it’s not."

The power of nonverbal feedback

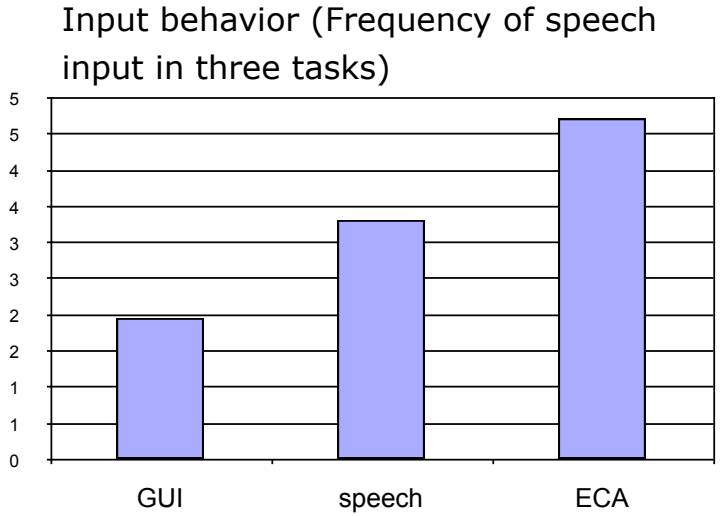
Ymir/Gandalf (Thorisson, 1996):

- Different kinds of feedback
  - Content-related: question answering, command execution
  - „Envelope": gaze and head movement for turn-taking/-giving and as attentional cues, coverbal beat gestures during speaking
  - Emotional: happy, puzzled face
- Fewer user repetitions and hesitations, better ratings of language capability of the system in content + envelope FB condition (Cassell & Thorisson, 1999).



Do agents induce more natural interactions?

- embodied interface agents **trigger user's speech input** – compared not only to text but also to speech based interfaces
- Users engage more frequently in **reciprocal communication** attempts such as correcting comments or resignation.



When interacting with agents users show behavior that merely is appropriate in human face-to-face interaction.

Problematic consequence: Human-like agents lead to **expectations** that can not yet be met.

(Krämer, 2005)

Social effects of an agent‘s presence

Agents cause social effects comparable to humans!

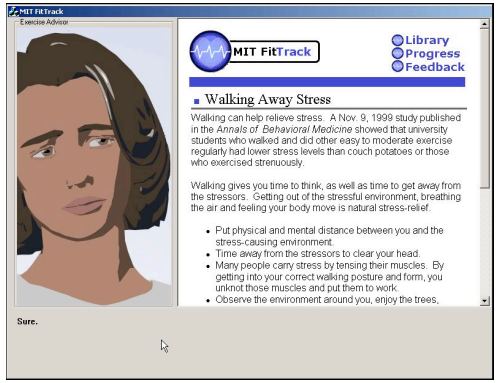
- Social presence of the virtual character
  - If confronted with an embodied interface agent, users try to present themselves in a more positive light (Sproull et al., 1996) ⇨ „impression management“
- Effects of social facilitation/social inhibition
  - Task-performance is inhibited by the social presence of a monitoring agent (Rickenberg & Reeves, 2000)
- Open questions
  - Long-term effects, or will 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?“*

Social agents

- Use human social interaction protocols in the interface
  - Small talk, mirroring to build rapport
  - Immediacy, flattery to show liking
  - Nonverbal social cues

ATR, Osaka

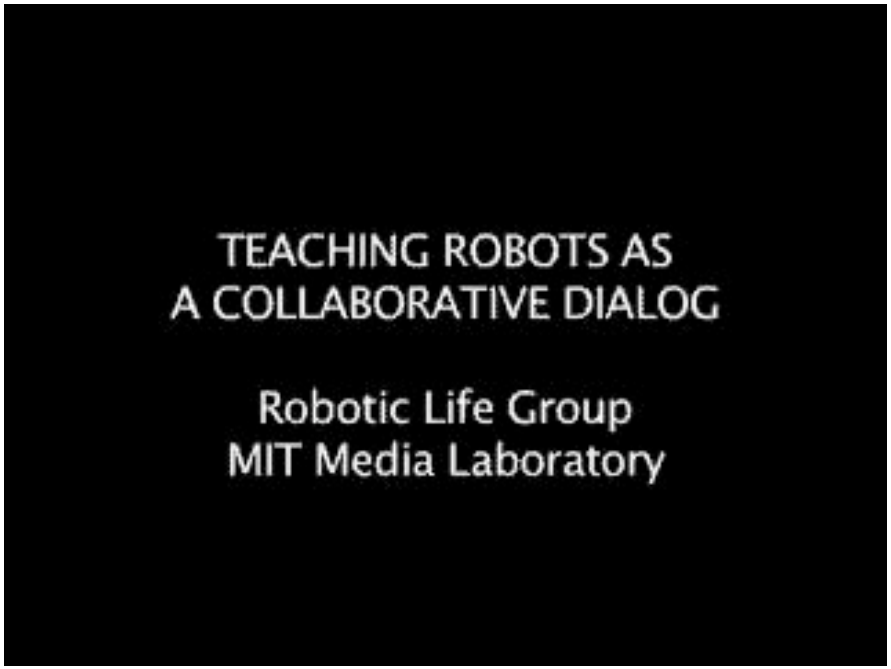
- Relational Agents:
  Computational artifacts designed to build and maintain long-term, socialemotional relationships with their users



Laura (Bickmore et al.)

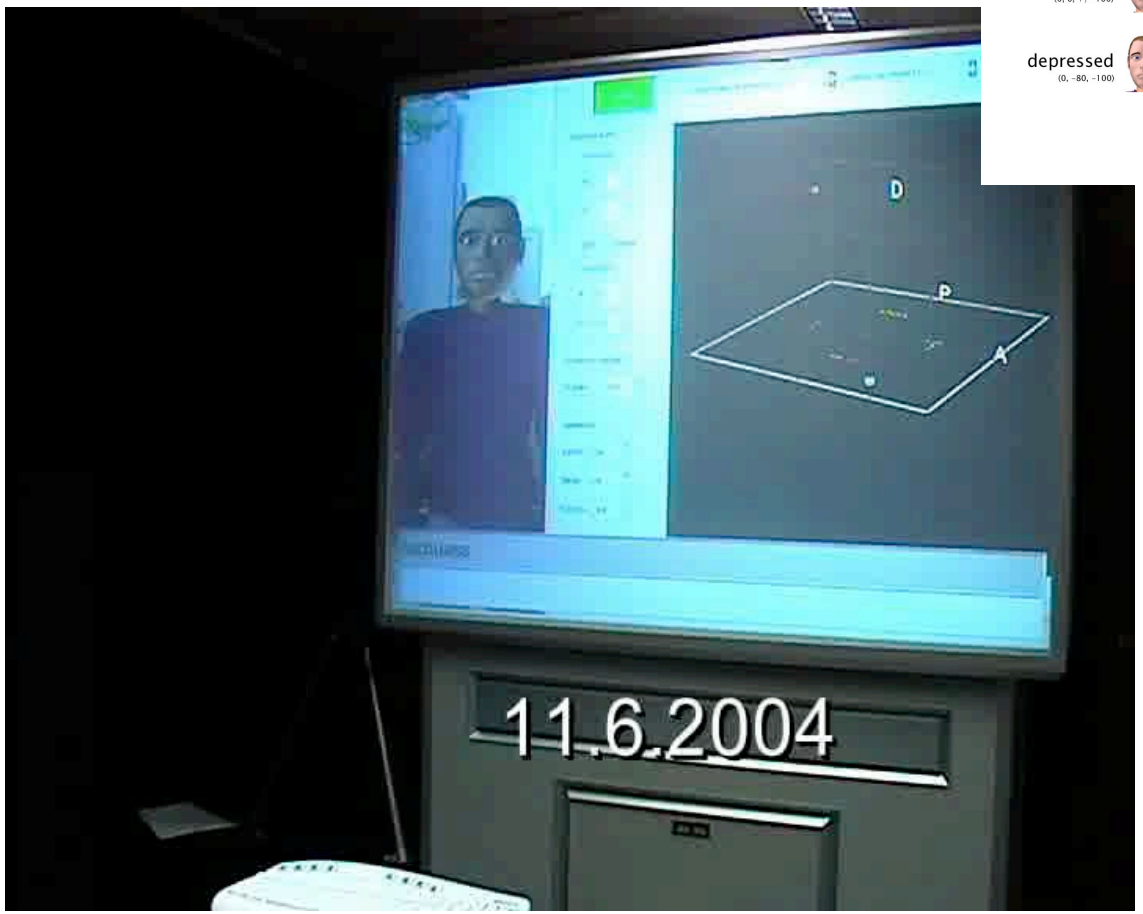
Social robots: Leonardo

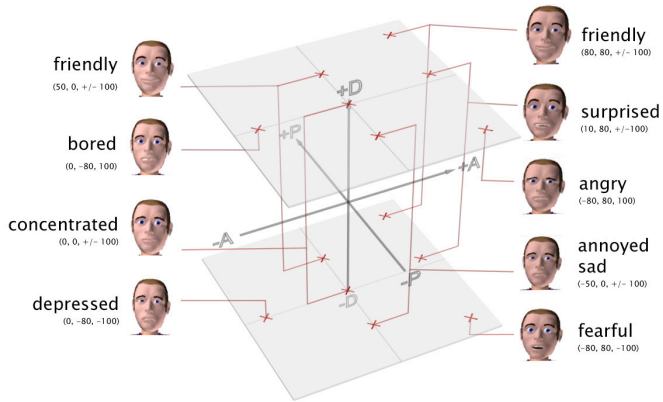
- Collaborative dialogue
- Modeling collaborative behavior by means of social cues
- Able to read same cues from human user and to provide mutual support back (intention recognition)
- Learning goal-directed actions through imitation and vision



C. Breazeal (MIT)

Emotional agents





Endow agents, their architectures and behavior, with emotions, feeling and personality

The final slide...

- Mensch-Maschine-Interaktion
  - Einführung & Historie
  - Kognitive Grundlagen: Perception, Memory, Attention, Reasoning
  - Interaction styles and technology
  - User-centered design and Usability Evaluation
  - Spoken Language Dialogue Interaction
  - Multimodal Interfaces
  - Agent-based Interfaces

- Klausur: 3.8.2006
  - genaue Zeit und Ort per Mail -> anmelden!!
  - Fragen zum Inhalt der Vorlesung