Human-Computer Interaction

Session 12
Agent-based interaction
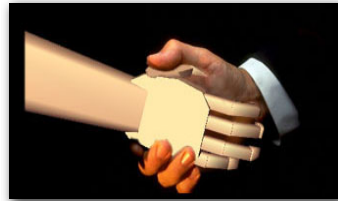
Evolution of user interfaces
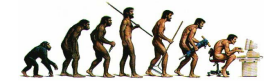
Year	Paradigm	Implementation
1950s	None	Switches, punched cards
1970s	Typewriter	Command-line interface
1980s	Desktop	Graphical UI (GUI), direct manipulation
1980s+	Spoken Natural Language	Speech recognition/synthesis, Natural language processing, dialogue systems
1990s+	Natural interaction	Perceptual, multimodal, interactive, conversational, tangible, adaptive
2000s+	Social interaction	Agent-based, anthropomorphic, social, emotional, affective, collaborative

---
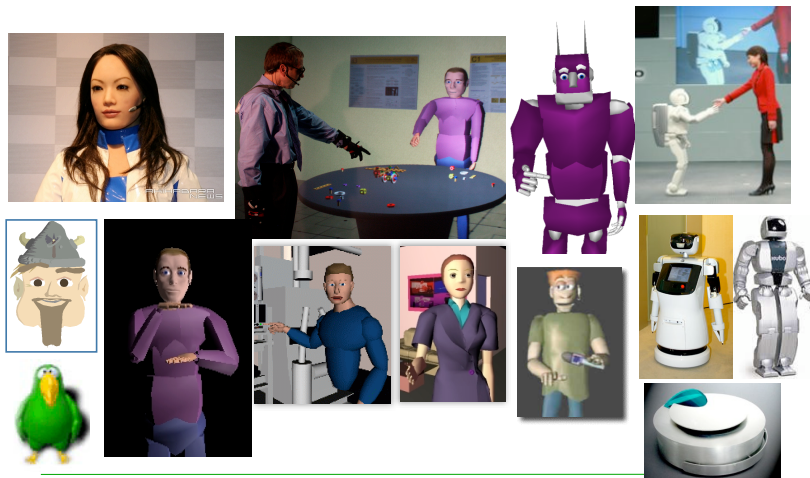
Interaction paradigms

☐ tools ➜ operate

☐ smart tools ➜ instruct

☐ interactive interlocutors ➜ converse

☐ companions ➜ collaborate

Agent-based interaction - ideas

☐ Build interactive software that...
- can communicate with other agents
- is autonomous, reactive, and proactive
- is context-aware and situated
- has specialized expert knowledge
- is personalized and user-adaptive

☐ In practice, the term agent applies 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

Terminology

☐ Agent
- computer system that observes and initiates actions in its environment, and is able to communicate with other individuals
- has a specific expertise and carries out specific tasks

"I'm interested in discussions on **agents**, **communityware**, **E-commerce**, and **HCI**."
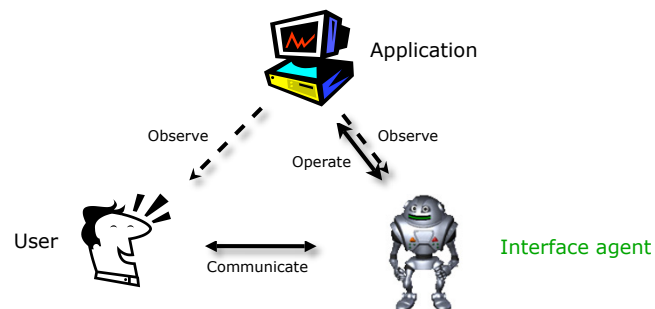
"OK, you might be interested in channel **VirSoc** (often discusses **communityware**)."

Some Time Later...

"They are discussing **agents** right now in channel **Hahvahd**."

Agents as direct interfaces

Application

Observe
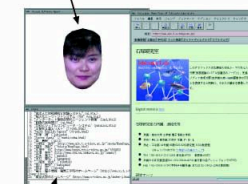
Observe

Operate

User

Communicate

Interface agent

☐ Agent mediates between the user and an application
- accepts the operating of the system
- communicates with and supports the user

Example: browser interface agent

(U. Tokyo)

Visual Software Agent
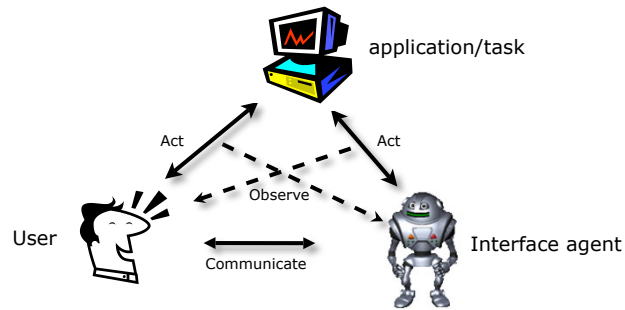*with rocking realistic face and speech dialog function.*

Netscape Navigator window

Anchor list
automatically extracted whenever new page is opened
<index_number, anchor_string, URL>

Collaborative interface agents

User and agent collaborate on a shared task
- User and agent can both take actions
- User and agent observe each other's actions
- User and agent communicate abouth their task and collaboration

---

„A buddy is better than a slave"

| Agent1 | We need to repair a connectivity problem between Mars and Saturn. Do a remote ping from Mars to Saturn. |
Agent2	can't. Saturn seems to be down. I'll take care of that first.
<Agent2 taking action>	
Agent2	Okay, Saturn's back up and the remote ping was successful.
Agent1	Good. Verify Mars' IP address for Saturn for me.
Agent2	The entry for Saturn was wrong, but I corrected it.
Agent1	Okay, good. We're done then.
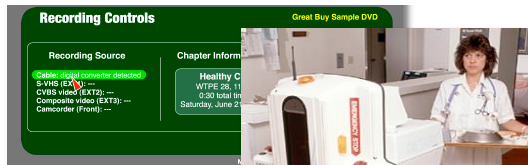
Interaction
- working on someone or something, possibly according to common rules, in order to achieve a personal goal
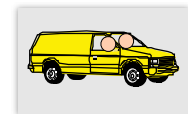

Collaboration
- working with others in order to achieve shared goals
  ➜ focus on working jointly
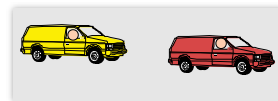
Cooperation
- working together with „somebuddy" to the same end
 ➜ focus on working & togetherness

So, not only division of labor...

Involves users actively in problem-solving
- leverage their skills
- steer solving process based on preferences or experiences
- increase user's trust, understanding, justifiability of solution

Example: Human-guided Search
(Klau et al. 2002)
- user can monitor, modify, or track back solutions
- user can apply, halt, or modify algorithms
- user can constrain and focus search
- improved performance, up to the best heuristic algorithms around
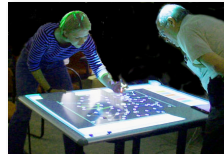
...but collaboration!

„Must design collaboration into systems from the start."

B. Grosz

Features of a multi-agent collaboration
- No master-slave relationship, but equality of partnership
- Agents have different beliefs, knowledge, and capabilities
- Agents share a goal and are committed to this goal
- Agents collaborate during both planning what to do and doing it
- Agents communicate to coordinate their collaboration

Collaboration

M. Bratman

Agent's intentions are crucial (Bratman 1987)
- commitment to action in order to achieve a goal
- constrain choices what else to intend
- provide context for re-planning upon failure
- guide means-ends-reasoning for plan refinement

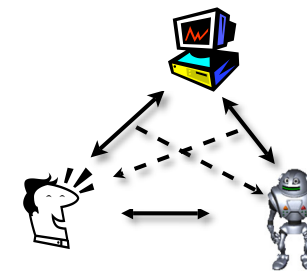Plans are mental states (Pollack 1990; Bratman 1990)
- not just knowing how to do an action (recipe)
- also having the intentions to do the actions entailed

Coordinating actions means coordinating minds

Collaborative agents

M. Bratman

Shared collaborative activity requires
1. mutual responsiveness
2. commitment to a joint activity
3. commitment to mutual support
4. meshing of subplans

SharedPlans formalism

- ☐ Formalizes how agents move from individual goals and intentions into collaborative, coordinated activity based on representations of the minds of the other agents:
  - ■ what is mutually believed and intended?
  - ■ what commitments have been taken by whom?

- ☐ Predominant model in multi-agent collaboration in A.I. and collaborative interfaces in HCI
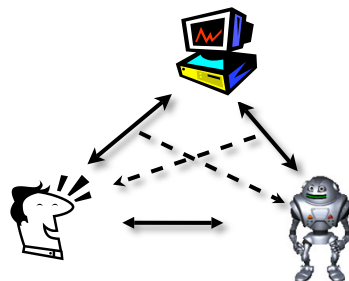

SharedPlans formalism

Collaboration starts by moving from one agent having a goal to a group having a SharedPlan to achieve it, via

- ■ explicit communication and conversational default rules
- ■ implicit plan recognition

An initially partial shared plan gets refined and augmented through reasoning, communicating, and group decision-making to become a full SharedPlan

- ■ each agent attributes to other(s) individual beliefs and intentions
- ■ each agent establishes mutual beliefs and intentions based on this and the context

---

Example: COLLAGEN

Charles Rich
Candace Sidner
Neal Lesh

Mixed-initiative problem solving assistant

- ■ employ SharedPlan formalism to manage what's called **collaborative discourse**
- ■ task-oriented spoken language dialogue

Collaborative discourse theory

intentional structure
- ☐ hierarchy of individual or shared goals and sub-goals (partial SharedPlan)

linguistic structure
- ☐ hierarchy of segments, each serving a purpose in the intentional structure

attentional structure
- ☐ context represented as focus stack of discourse segments

Intentional

goals, recipes, plans

focus spaces, focus stack

segments, lexical items
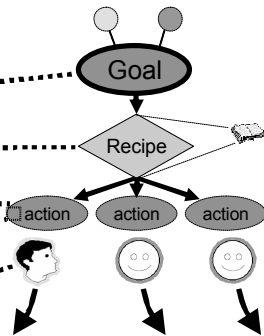
Attentional   *Linguistic*
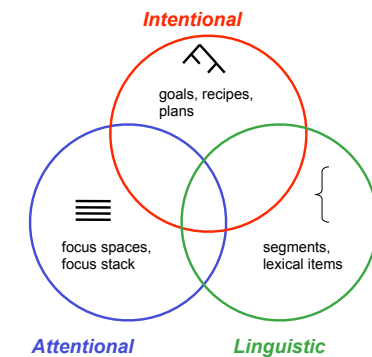
Collaborative discourse

Interactions identified by purpose:

- directly achieve current goal
- identify recipe to be used for goal
- achieve step in recipe for goal
- specify parameter of step/goal
- identify agent to perform step/goal

Hierarchy used to track content
and context of discourse …

Collagen

Example: Daimond Help

X-TECH	Washer/Dryer	ZONI	Thermostat
X-TECH	Refrigerator/Oven	TECH-E	Microwave
TECH-E	Home Theater	TECH-E	DVD Recorder
ZONI	Security System	ZONI	Lawn Sprinklers
MERL	Simple Demo		
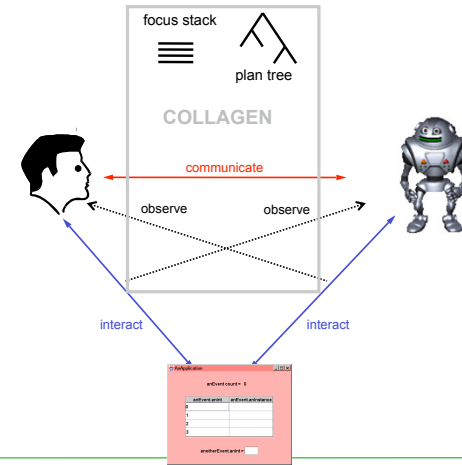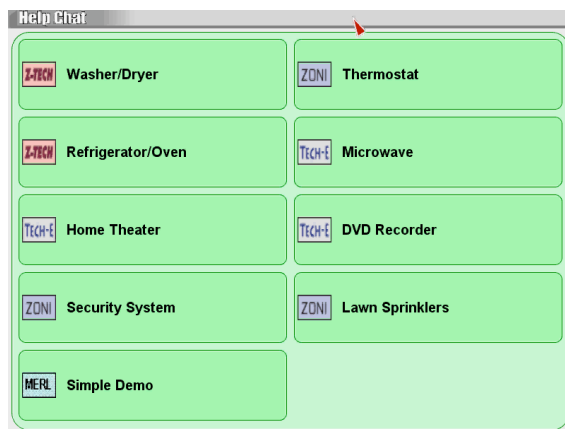
Pro & contra of agent-based interfaces

BEN SHNEIDERMAN
(U. Maryland)

vs.

PATTIE MAES
(MIT Media Lab)
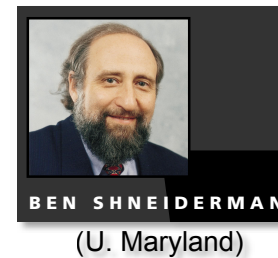
Ben Shneiderman and Pattie Maes debated these issues and more at
both IUI 97 (Intelligent User Interfaces conference - Jan 1997) and
CHI 97 (March 1997)

Ben Shneiderman

„Users should **comprehend** the display, feel in **control**, be able to **predict** the system, take **responsibility** for their actions"

„**Responsibility** will be the central issue in this debate."

„**Direct manipulation**: rapid, reversible, incremental, point & click, immediate feedback, reduces error, encourages exploration"

„Future is moving in the direction of information **visualization**"

„**Overview** is most important, giving users a sense of context."

„Anthropomorphic or social interface is **not** to be the future of computing."

Pattie Maes

„Agents are **personalized**, **proactive**, **long**-**lived**, **adaptive** to user, acts on user's behalf based on knowledge of user preferences"

„Necessary because environment becomes complex, users become naive, number of tasks and issues increase"

„Agents are **no alternatives to direct manipulation**, nor are they necessarily personified or deal with NL interaction. You still need a well-designed interface when incorporating agents in an application. However, some task I may just not do myself."

„Using an agent **doesn't imply giving up all control**, just over the details and that saves me a lot of time."

„The true **challenge** lies in **designing the right user-agent interface.**"

Shneiderman

„Speech is important for niches but will **not** be a generally usable tool, and it **degrades** your problem solving performance."

„Anthropomorphic representation **misleads** designer, **deceives** users, increases **anxiety** about computer usage, **interferes** with predictability, **reduces** user control, **undermines** users' responsibility."

„Users want to have the feeling that **they** did the job-not some magical agent."

„**human-to-human interaction is not a good model** for the design of user interfaces."

„Get past the argumentation about a system being more friendly or more natural or intuitive, focus on real user performance and real tasks. Do your **scientific evaluation**."

Maes

„A good user-agent interface takes care of two issues: **understanding** (of the agent) and user's felt control over tasks but its **possible delegation** to the agent."

„Most successful interfaces are the ones where the agents are pretty much **invisible**."

„Ben focuses on professional users and well-structured task domains and well-organized information domains. We are dealing with **untrained** end users and **ill-structured** and **dynamic** information domain."
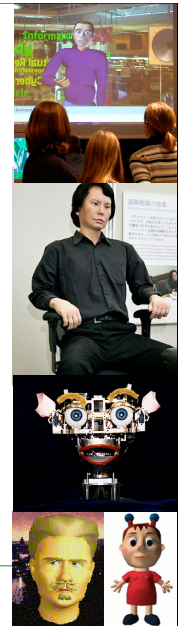
„Users **do not** always want to have **all of control**."

Embodied conversational agents

Face-to-face conversation as user-agent interface

Terminology

Embodied Agent

- equipped with a human-like body
- employs body for action and communication
- aspires human-like use of modalities and communication protocols of face-to-face conversation


Terminology

Embodied Conversational Agent

- Recognize and interpret verbal and nonverbal input behavior
- Generate output behavior
- Process the multiple functions of conversational behavior
- Take an active role in dialogue

„Computer interfaces that hold up their end of conversational, 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"

(Justine Cassell 2000)
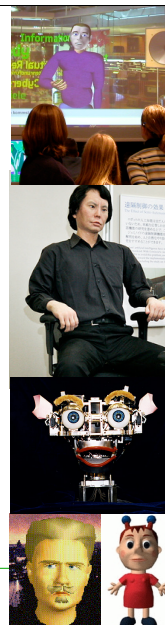
Embodied agent interfaces - motives

Interaction is more intuitive
- familiar communication and interaction strategies apply

Tasks appears less complex in a team
- expertise and proactivity of the agent supports the user (e.g., expert critics, subtask completion, coordination)
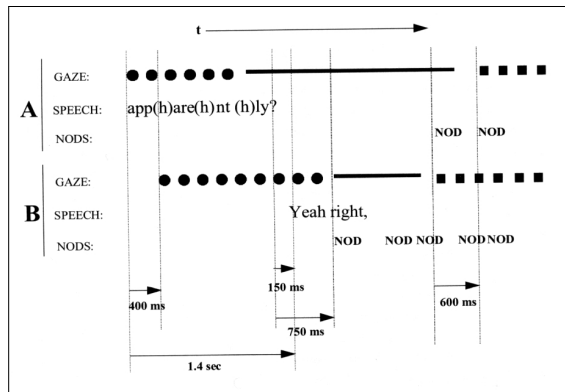
Metaphor of a mediator becomes tangible
- „somebody" is there, with me, and helps me out (a persona)

Motivational and social factors
- interacting with „somebody" is more entertaining and motivating, entails socio-affective effects
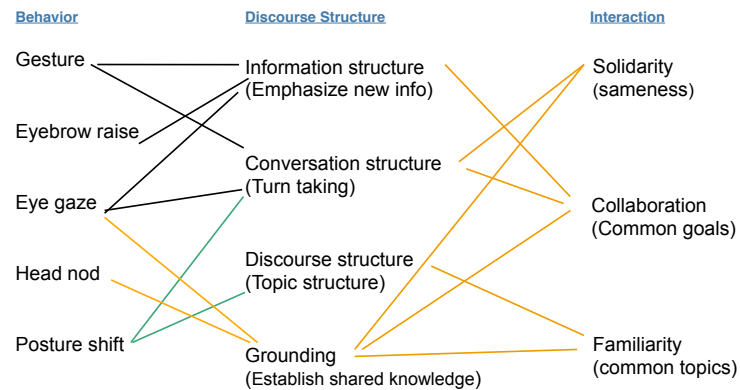
Conversational behavior

Conversational behavior

What's needed:

- use multiple, finely synchronized modalities to pursue interactional and propositional goals in parallel
- fulfill conversational functions (e.g., turn taking, turn keeping, feedback, emphasize) realized by communicative behaviors
  - a behavior may convey several function's; a function may be realizable by different sets of behaviors

Example:
Turn-taking

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

Functions of nonverbal behavior

Behavior — Discourse Structure — Interaction

Gesture

Eyebrow raise

Eye gaze

Head nod

Posture shift

Information structure (Emphasize new info)

Conversation structure (Turn taking)

Discourse structure (Topic structure)

Grounding (Establish shared knowledge)

Solidarity (sameness)

Collaboration (Common goals)

Familiarity (common topics)

**Justine Cassell
ArticuLab**

Why is it so hard to build an ECA?

Conversational behavior theory

- Model of cf and cb's and their mappings needed for 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

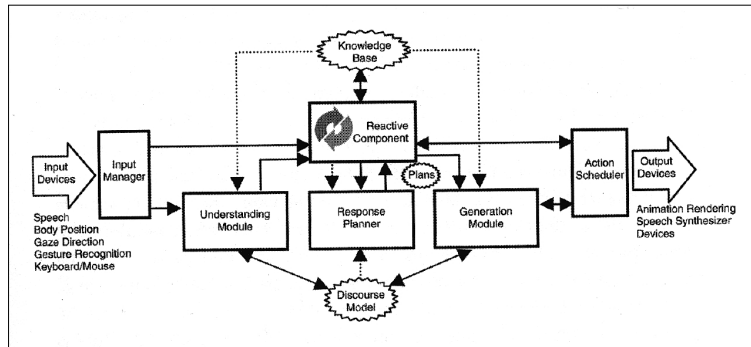Multistep deliberation, parallelism, modularity

- Input understanding, response/dialogue planning, and output generation must run fast, parallel

Timing & efficiency

- Different threads of communication must be handled at different timescales

Output synchrony

ECA architectures

System examples

Animated Conversation (1994)

System examples

REA - the real estate agent (MIT Media Lab, 2000)

(Cassell et al.,
1999, 2000)

System examples

Cultural training (ICT/ISI, L.A.)

VIRTUAL HUMANS
for
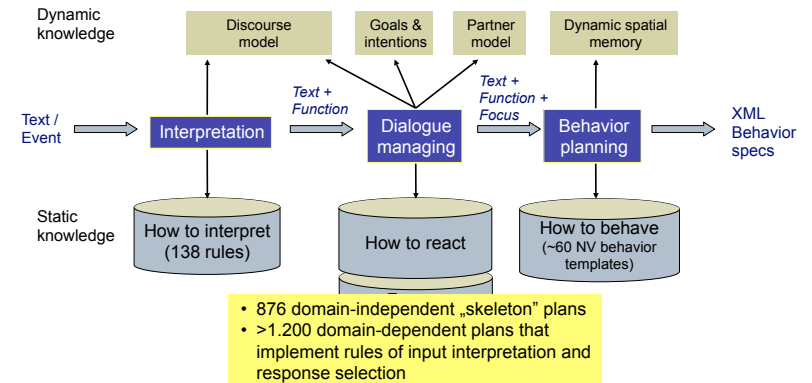Training Stability and
Support Operations

System examples

Max (AG-WBS, Uni Bielefeld)
Here, as conversational museum guide in the HNF

Plan-based dialogue system

Dynamic knowledge

Discourse model | Goals & intentions | Partner model | Dynamic spatial memory

Text / Event → Interpretation → *Text + Function* → Dialogue managing → *Text + Function + Focus* → Behavior planning → XML Behavior specs

Static knowledge

How to interpret (138 rules) | How to react | How to behave (~60 NV behavior templates)

- 876 domain-independent „skeleton" plans
- >1.200 domain-dependent plans that implement rules of input interpretation and response selection

---

Effects of ECAs?

Virtual faces draw attention (Dehn & van Mulken, 2000)

Interaction tends to be more entertaining (Takeuchi & Naito, 1995; Koda & Maes, 1996; van Mulken et al., 1998, Krämer et al., 2002)

Acceptance is higher (Hubona & Blanton, 1996; Ahern, 1993)

Increased intelligence, trustworthiness, believability (Sproull et al. 1996; Walker, Sproull & Subramani, 1994; Rickenberg & Reeves, 2000)

User are more inclined to delegate tasks to the system (Milewski & Lewis, 1997)

Natural language & reciprocal communication (Krämer, 2005)

Social dialogue, politeness (Kopp et al. 2005, Bickmore 2003)
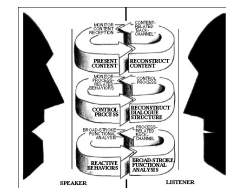

Effects of ECAs?

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
- ☐ Emotional: happy, puzzled face

Fewer user repetitions and hesitations, better ratings of language capability of the system in content + envelope FB condition

Positive effects of emotional feedback on user evaluation (V.d. Pütten et al. 2009)
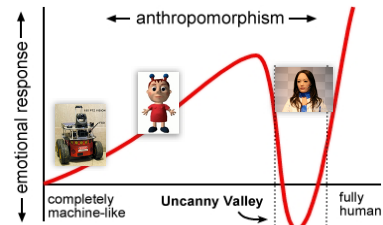

Ymir/Gandalf
(Thorisson, 1996)
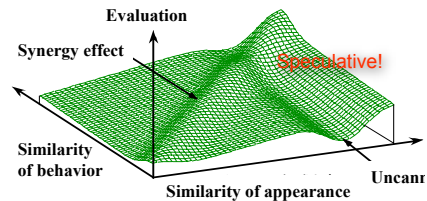
(Cassell & Thorisson, 1999)

Effects of ECAs?

Human-like appearance
- impact on social evaluation
- similarity with self
- high realism raises high expectations

Hypothetical „uncanny valley"
(Mori 1970; Ishiguro (2005)

Embodied collaborators

Robotic „partners"?

(Breazeal et al. 2004)

Problem: today robots don't interact with people as people
- not aware of other's goals and intentions
- don't adjust their behavior to help us
- no joint attention, no spatial or mental perspective-taking
- don't know what's hard to access or important for the human
- don't communicate to establish shared beliefs, coordinate, and demonstrate commitment
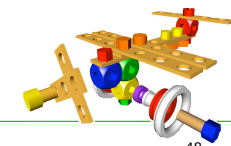- don't live up to the social models that humans use to understand and predict behavior

---

Max as a cooperation partner

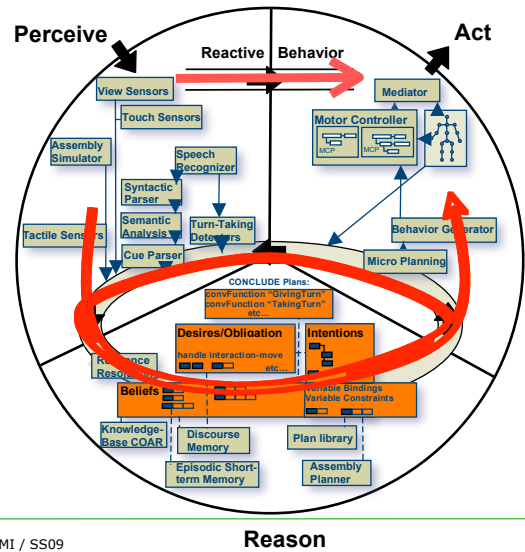Situated Artificial Communicators
SFB 360

- Face-to-face interaction with the embodied agent *Max* in an immersive Virtual Reality environment

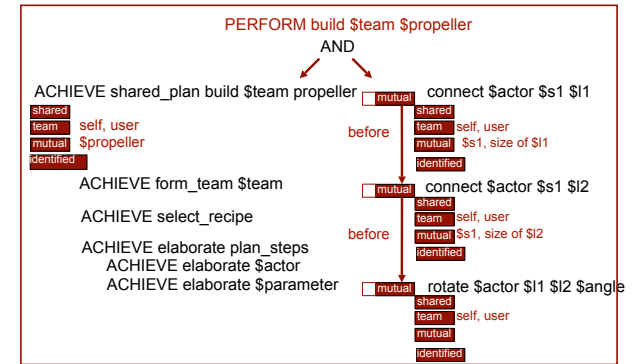- Study communication in a cooperative construction task
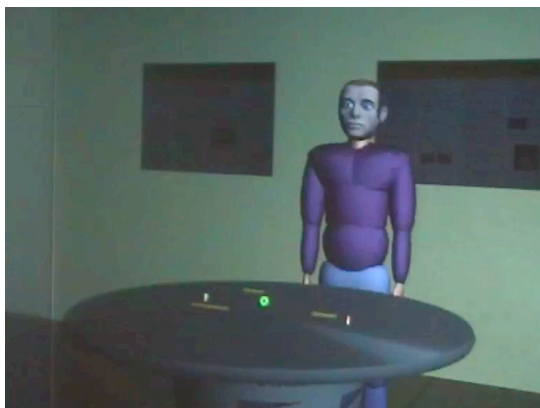

Nadine Pfeiffer-Leßman

Perceive **Act**

Reactive Behavior

Reason

Shared plans, mutual beliefs, intention for the group to succeed

PERFORM build $team $propeller
AND

ACHIEVE shared_plan build $team propeller — mutual — connect $actor $s1 $l1
shared
team — self, user
mutual — $propeller
identified

before

shared
team — self, user
mutual — $s1, size of $l1
identified

ACHIEVE form_team $team — mutual — connect $actor $s1 $l2
shared
team — self, user
mutual — $s1, size of $l2
identified

ACHIEVE select_recipe

before

ACHIEVE elaborate plan_steps
ACHIEVE elaborate $actor
ACHIEVE elaborate $parameter — mutual — rotate $actor $l1 $l2 $angle
shared
team — self, user
mutual
identified

Example

AG WBS, Univ. Bielefeld

Next session:
Social aspects of embodied HCI