Universität Bielefeld

5 Kommunizierende Agenten

15. Vorlesung: Max, unser Agent in der virtuellen Welt

Methoden der Künstlichen Intelligenz

Ipke Wachsmuth WS 2008/2009

Fragen der Vorlesung auch:

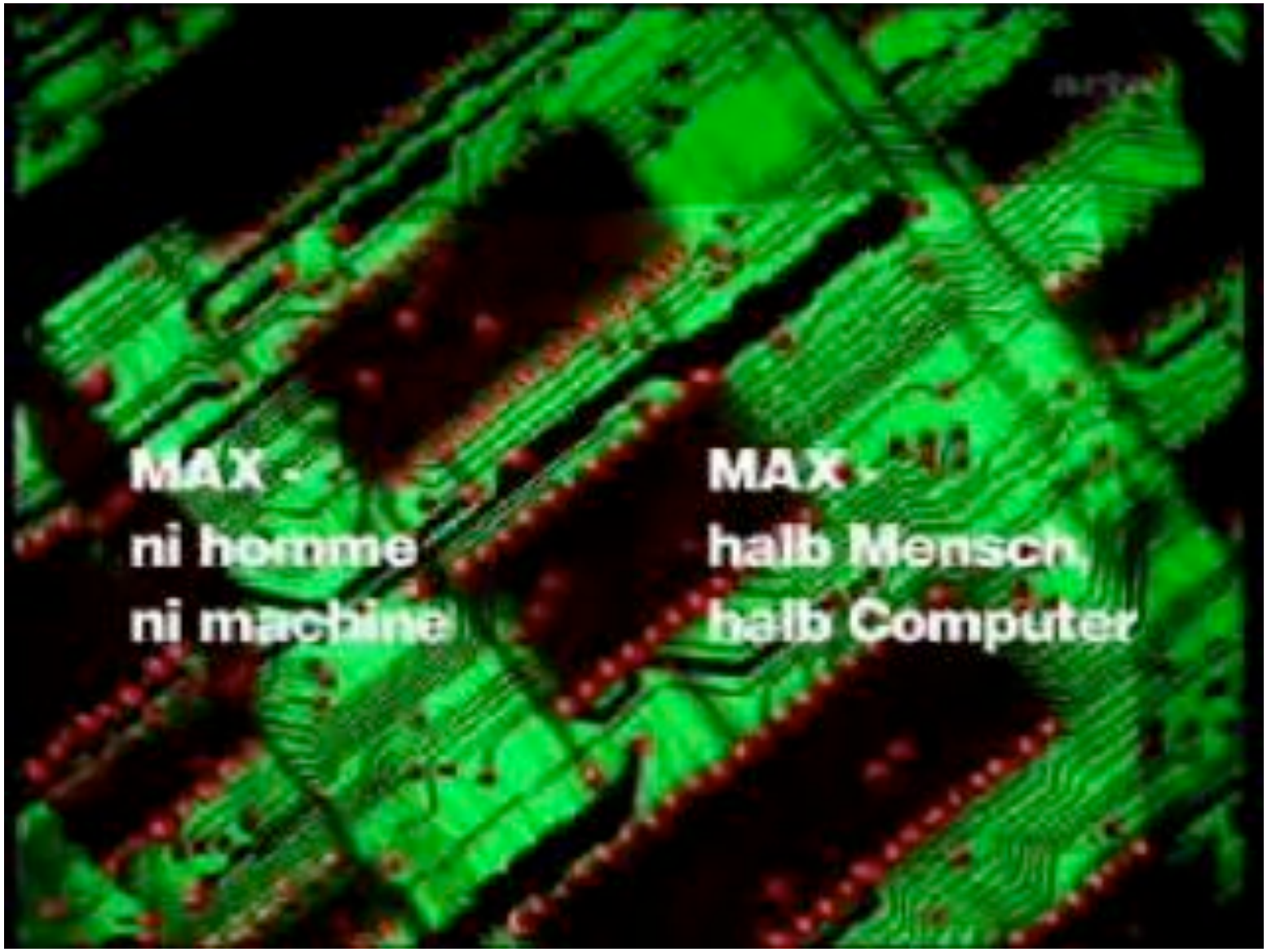
Werden / können / sollen KI-Systeme menschenähnlich sein?

Androiden

Humanoide Roboter

Virtuelle Menschen

Unser Agent in der virtuellen Welt



Kommunikationsforschung in der KI

"Learning to generate is learning to understand"

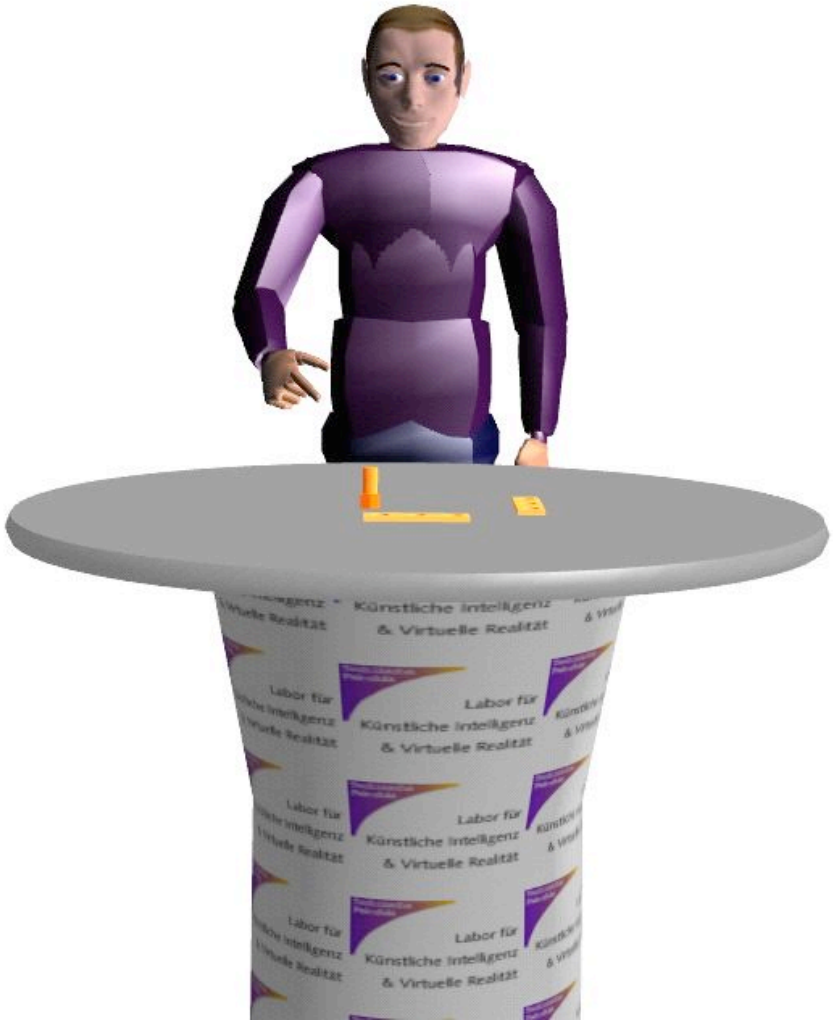
- Erklärungsmodelle in der Form von Verhaltens- und Prozess-Simulationen
- kognitive und soziale Faktoren von Kommunikation besser verstehen

Neue Formen der Mensch-Technik-Schnittstelle

- Humanoide Agenten sehr interessant in der Mensch-Maschine-Interaktion
- kommunizierende Agenten könnten zu „Partnern“ des Menschen werden

Agent MAX

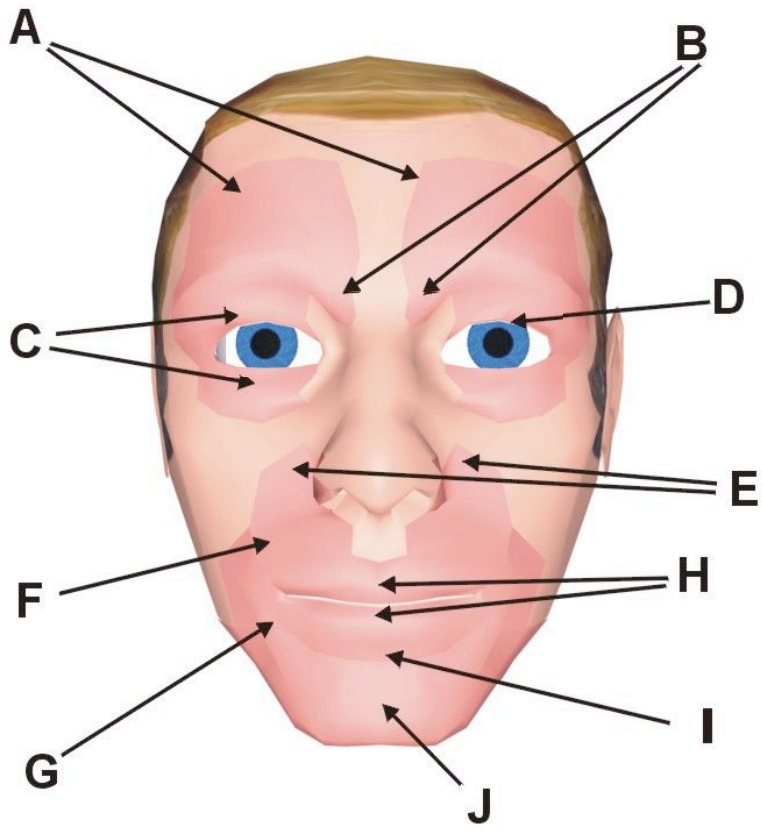
Ein künstlicher Kommunikator, situiert in virtueller Realität



Erforschung von Grundlagen kommunikativer Intelligenz im „face-to-face“-Dialog:

- PHYSIS – das Körpersystem (speziell Gestik, Mimik)
- KOGNITION – das Wissens-system (speziell Sprache)
- EMOTION – das Bewertungssystem

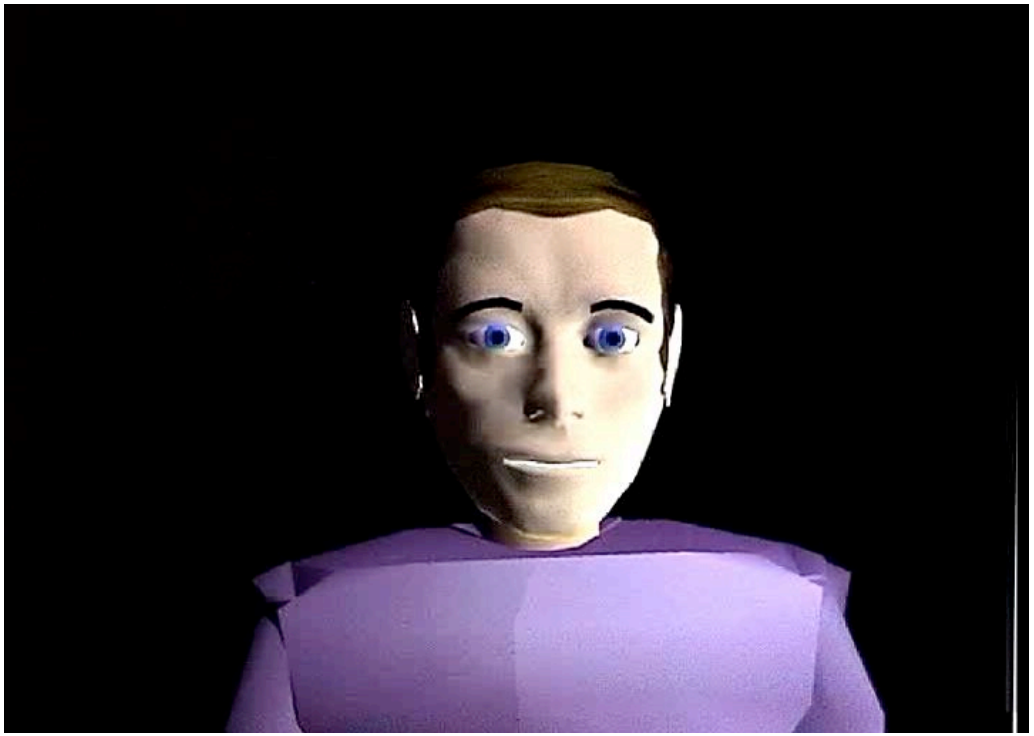
Simulierte Gesichtsmuskeleffekte



Studentisches Projekt
(Körber-Studienpreis!)

A	Stirnmuskel
B	Augenbrauenrunzler
C	Augenringmuskel
D	Augenlidmuskel
E	Heber d. Oberlippe und d. Nasenflügels
F	Jochbeinmuskel / Mundwinkelheber
G	Mundwinkelherabzieher
H	Ringmuskel des Mundes
I, J	Unterlippenherabzieher, Unterkiefer

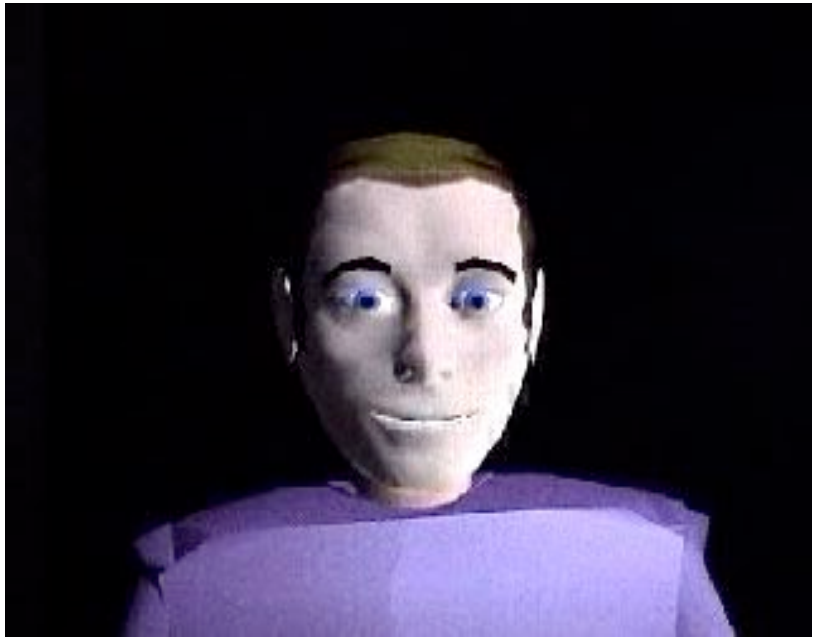
Animation: Gesicht und Körper



- Gesichtsanimation mit sog. Aktionseinheiten
- Berührungssensoren
- simuliertes Emotionssystem

- Max hat ein bewegliches Skelett mit 103 Freiheitsgraden
- 53 für den Körper und 25 für jede Hand
- synchronisierte Bewegungen

Sprechen und Sprechmimik



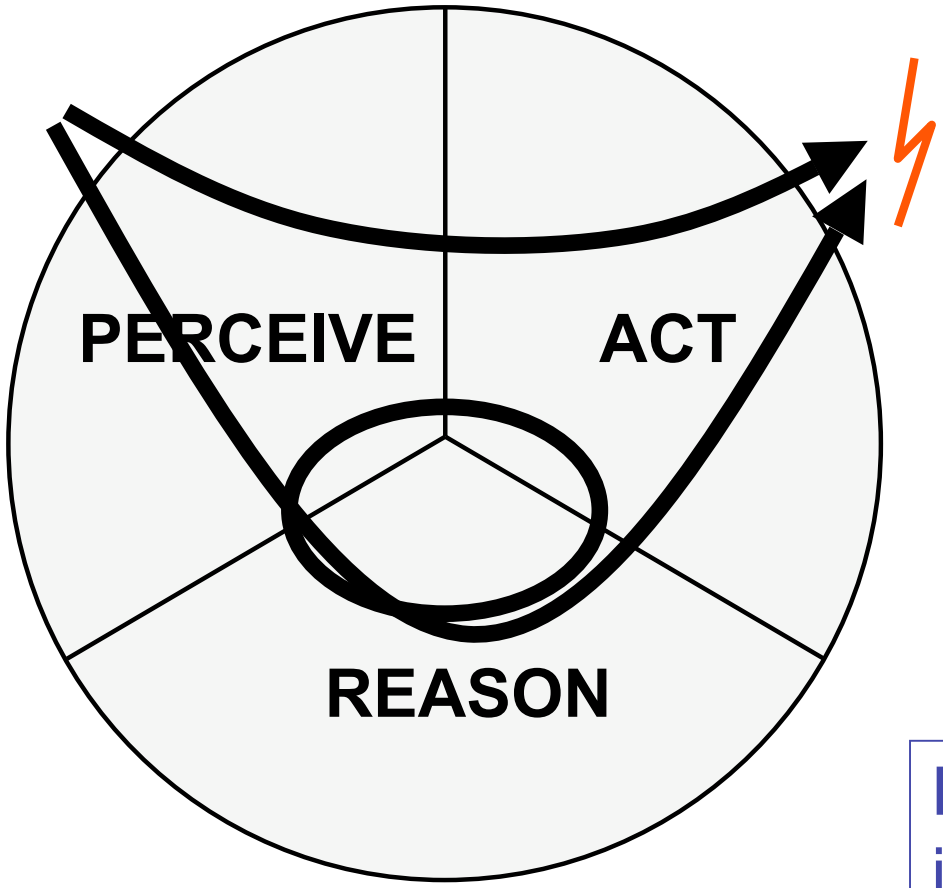
Text-to-Speech:

TXT2PHO (IKP Uni Bonn), MBROLA
Phonemtranskription ist Grundlage für die automatische Generierung von sog. Visemen für die Sprechmimik

(Concept-to-Speech: in Arbeit)

- eine Gesichtsstellung (Visem) für *M, P, B*
- eine Gesichtsstellung (Visem) für *N, L, T, D*
- eine Gesichtsstellung (Visem) für *F, V*
- eine Gesichtsstellung (Visem) für *K, G*
- und Viseme für die verschiedenen Vokale

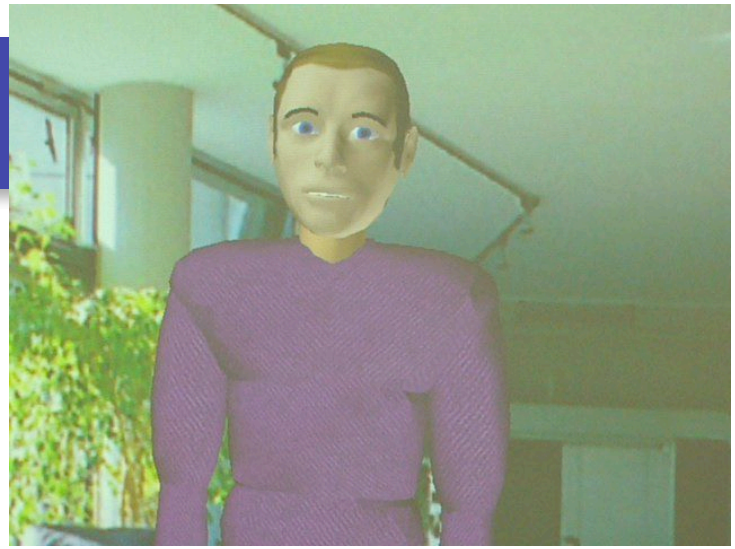
PERCEIVE – REASON – ACT

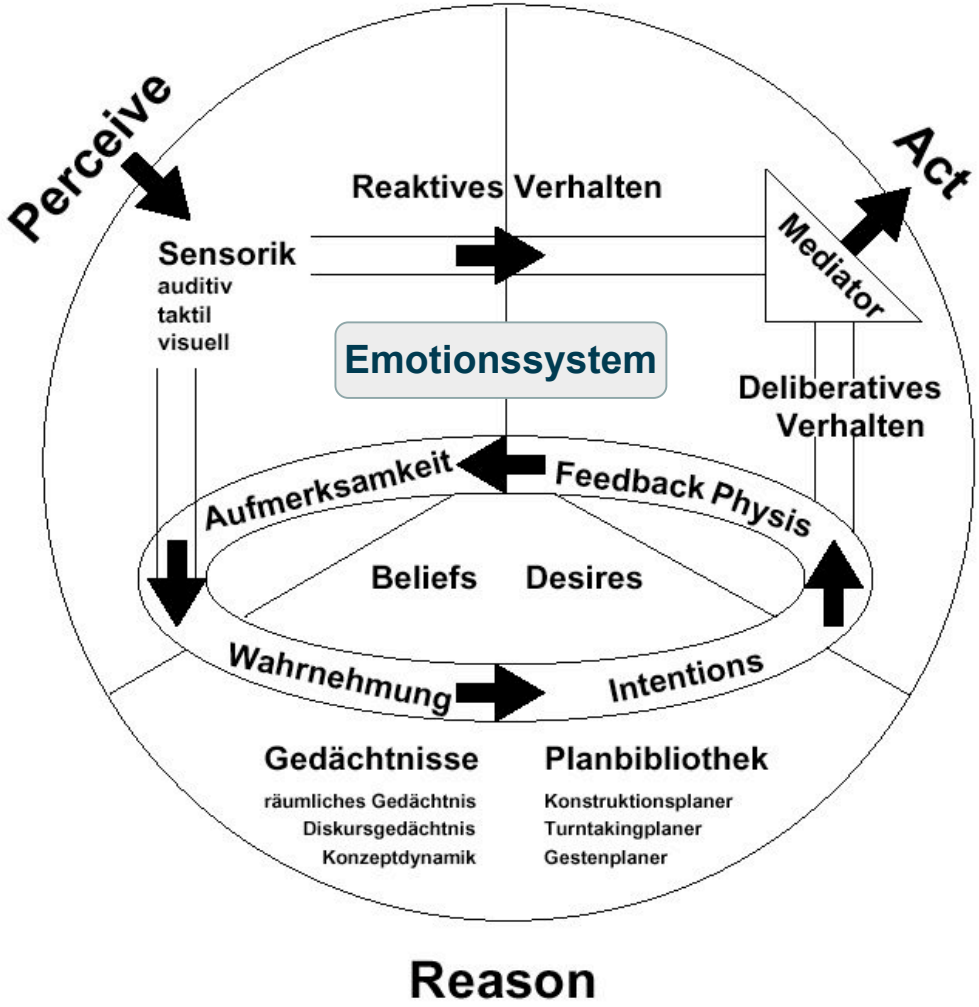


- Reaktives und deliberatives Verhalten kann nebenläufig sein und wettbewerben!
- Eine impulsive Reaktion kann durch kontrolliertes Handeln überschrieben werden, und umgekehrt.

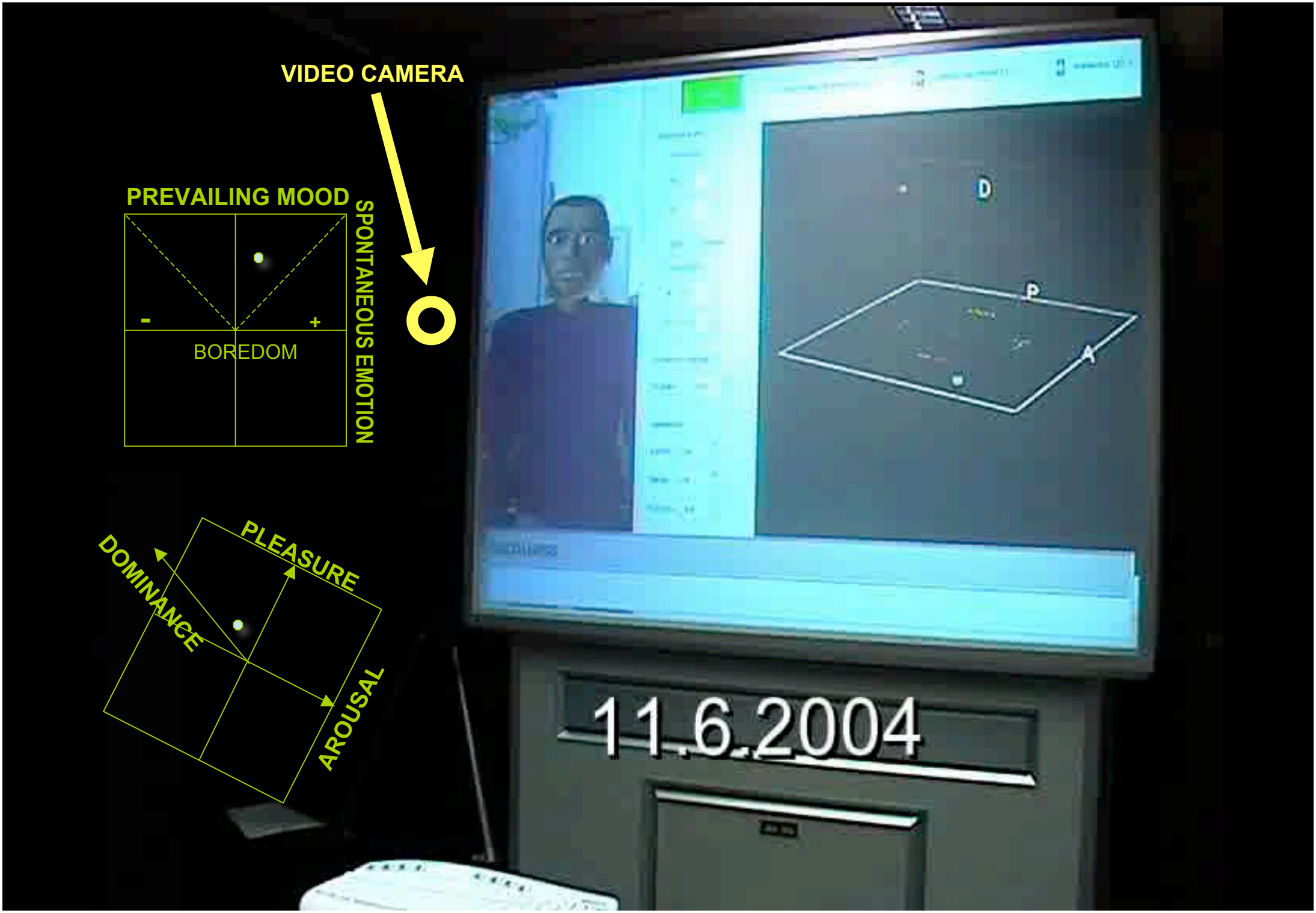
Deliberatives Verhalten benötigt interne Repräsentationen!

Kognitive Architektur

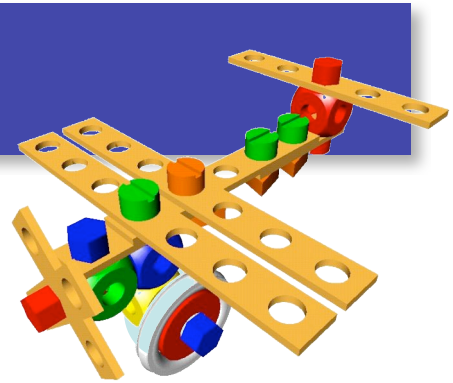




- nebenläufiges Wahrnehmen, Schlussfolgern und Handeln
- paralleles reaktives und deliberatives System
- BDI-Kern mit dynamischen eigenständigen Planern
- Physis, Multimodalität und Emotion: u.a. Zielbewertung


VIDEO CAMERA
PREVAILING MOOD
SPONTANEOUS EMOTION
-
+
BOREDOM
PLEASURE
DOMINANCE
AROUSAL
D
P
A
11.6.2004

Kooperativer Dialog



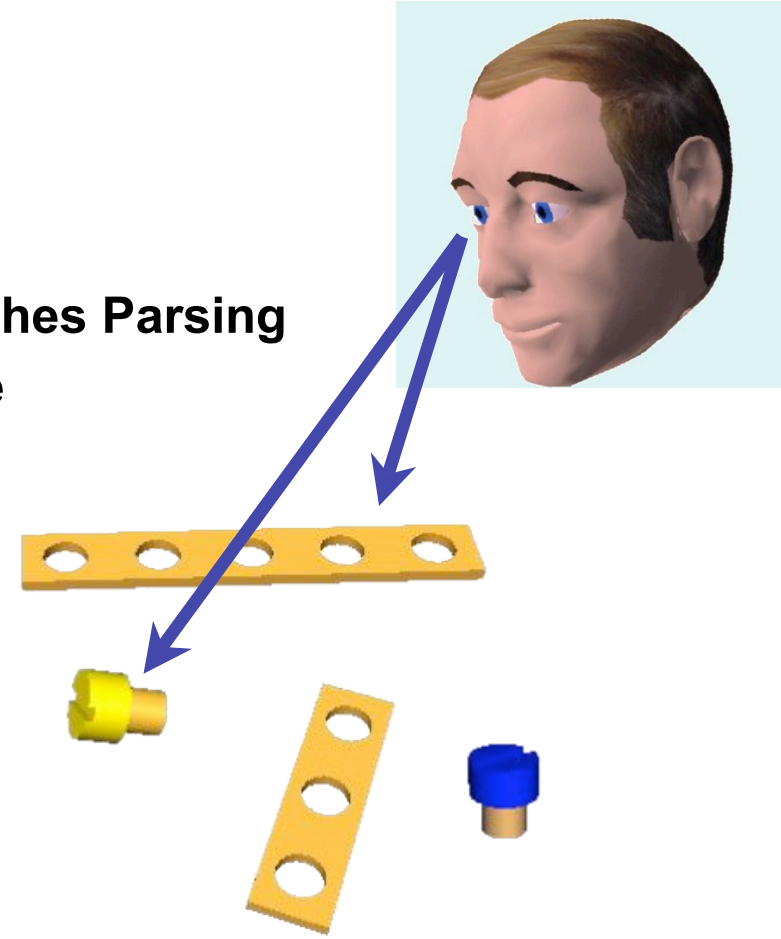
Zwei Partner bauen ein Modellflugzeug zusammen

Instrukteur (I) ist ein Mensch, Konstrukteur (K) ein künstlicher Agent

I und K führen einen „kooperativen Dialog“

Sprache verstehen

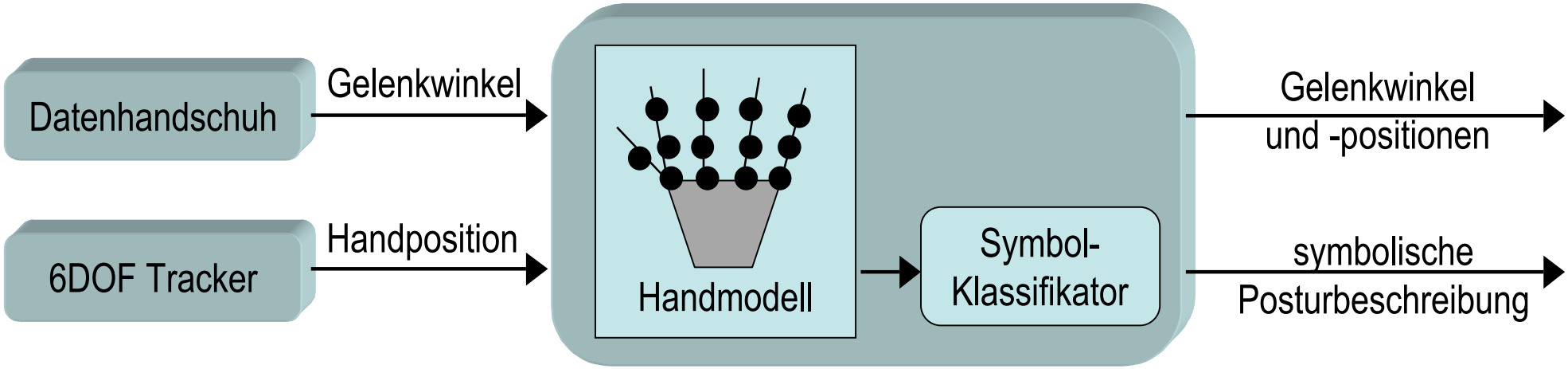
- **Spracherkennung**
- **Syntaktisch-semantisches Parsing**
- **Referenz auf die Szene**

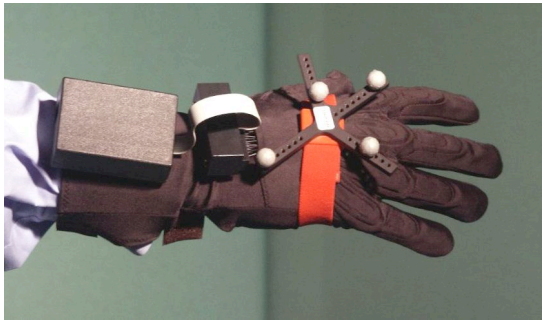


```
steck     BEFEHL      VERBINDEN
die       ARTIKEL
gelbe     FARBE       GELB
Schraube  OBJEKTTYP   SCHRAUBE
in        PRAEP       IN
die       ARTIKEL
lange     GROESSE     GROSS
Leiste    OBJEKTTYP   LEISTE
```

```
(select x (OBJEKTTYP(x)= SCHRAUBE und FARBE(x)= GELB))
(select y (OBJEKTTYP(y)= LEISTE und GROESSE(y)=GROSS))
```

Gesten verstehen

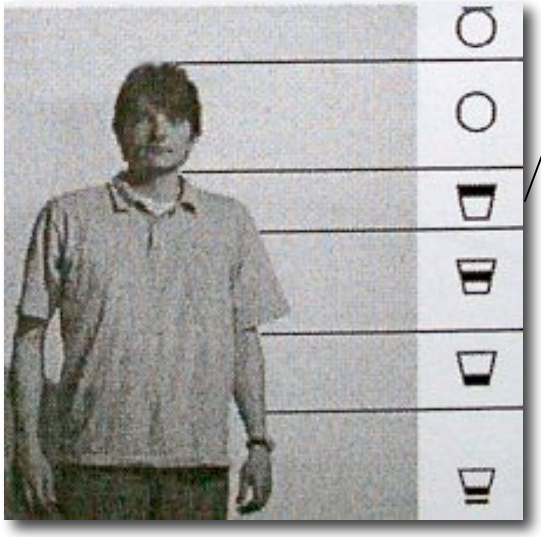




Symbolische Beschreibung der Gestenform (HamNoSys)

(„Hamburger Notations-System“ des Instituts für Deutsche Gebärdensprache, Hamburg)

Einige HamNoSys-Symbole



Symbol	ASCII-Darstellung	Beschreibung
	`BSifinger`	Zeigefinger gestreckt
	`EFinA`	nach vorn gestreckt
	`PalmL`	Handfläche nach links
	`LocShoulder`	in Schulterhöhe
	`LocStretched`	ganz ausgestreckt
	`MoveA`	Handvorwärtsbewegung
	`MoveR`	Handrechtsbewegung
<usw.>	...	...
( )	( )	parallel ausgeführt
[ ]	[ ]	in Sequenz ausgeführt

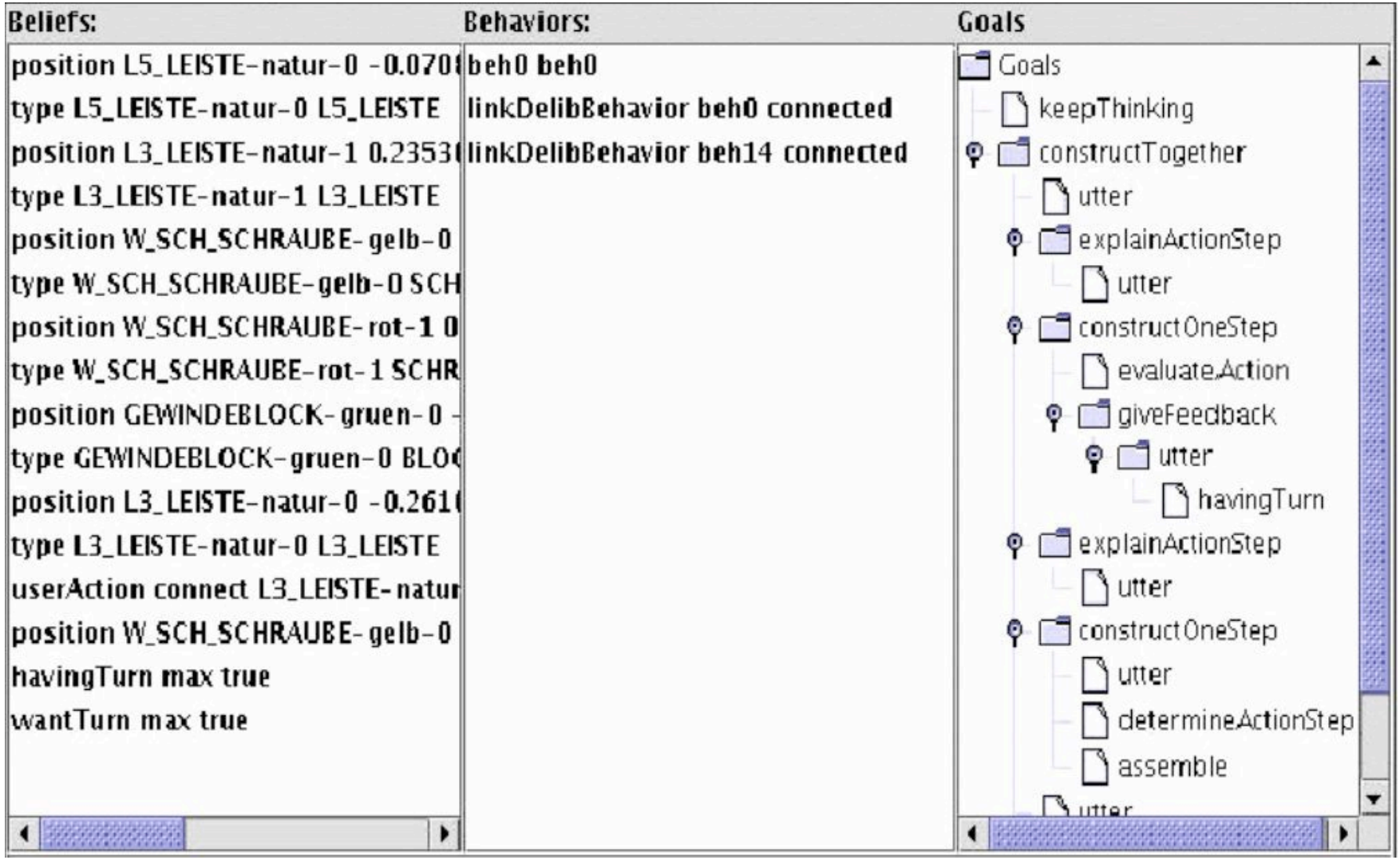
Gestik und Sprache äußern

MURML: XML-basierte Markup Language für „Multimodal Utterance Representations“

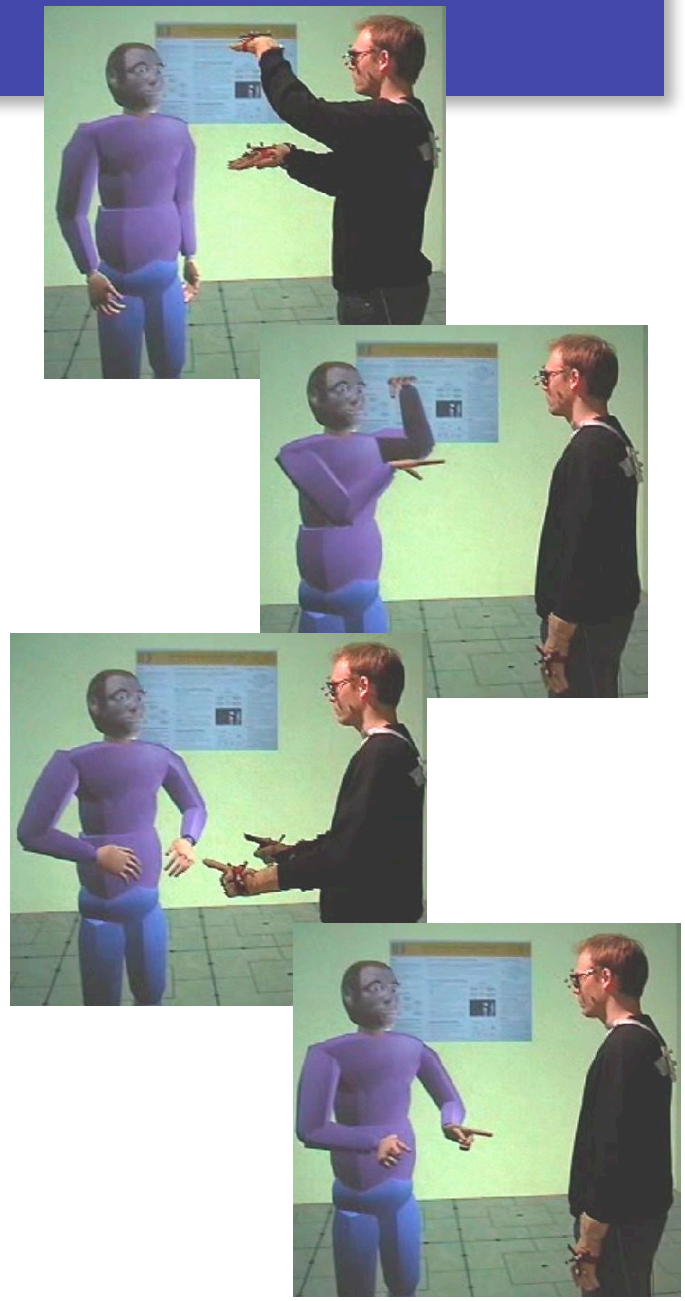
```
<utterance>
  <specification>
    Und jetzt nimm <time id="t1"/> diese Leiste
    <time id="t2“ chunkborder="true"/>
    und mach sie <time id="t3“/> so gross. <time id="t4"/>
  </specification>
  <behaviorspec id="gesture_1">
   <gesture>
    <affiliate onset="t1" end="t2"/>
    <constraints>
     <parallel>
      <static slot="HandShape" value="BSifinger"/>
      <static slot="ExtFingerOrientation"
            value="$object_loc_1“ mode="pointTo"/>
      <static slot="GazeDirection" value="$object_loc_1"
            mode="pointTo"/>
     </parallel>
```

Kommunizieren mit Agent Max

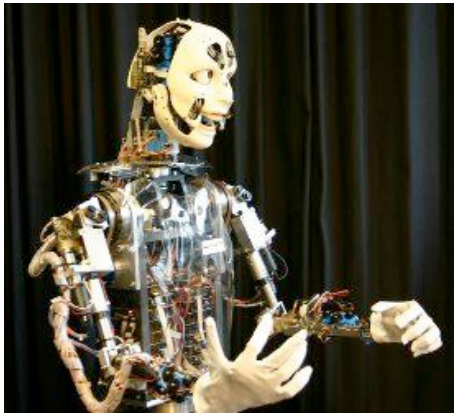
Interne Repräsentationen



Lernen durch Imitieren



- Kann Max neue Gesten durch Imitation lernen?
- Können wir dies auf humanoide Roboter übertragen?



  - BARTHOC (SFB)
  - Asimo (CoR-Lab)

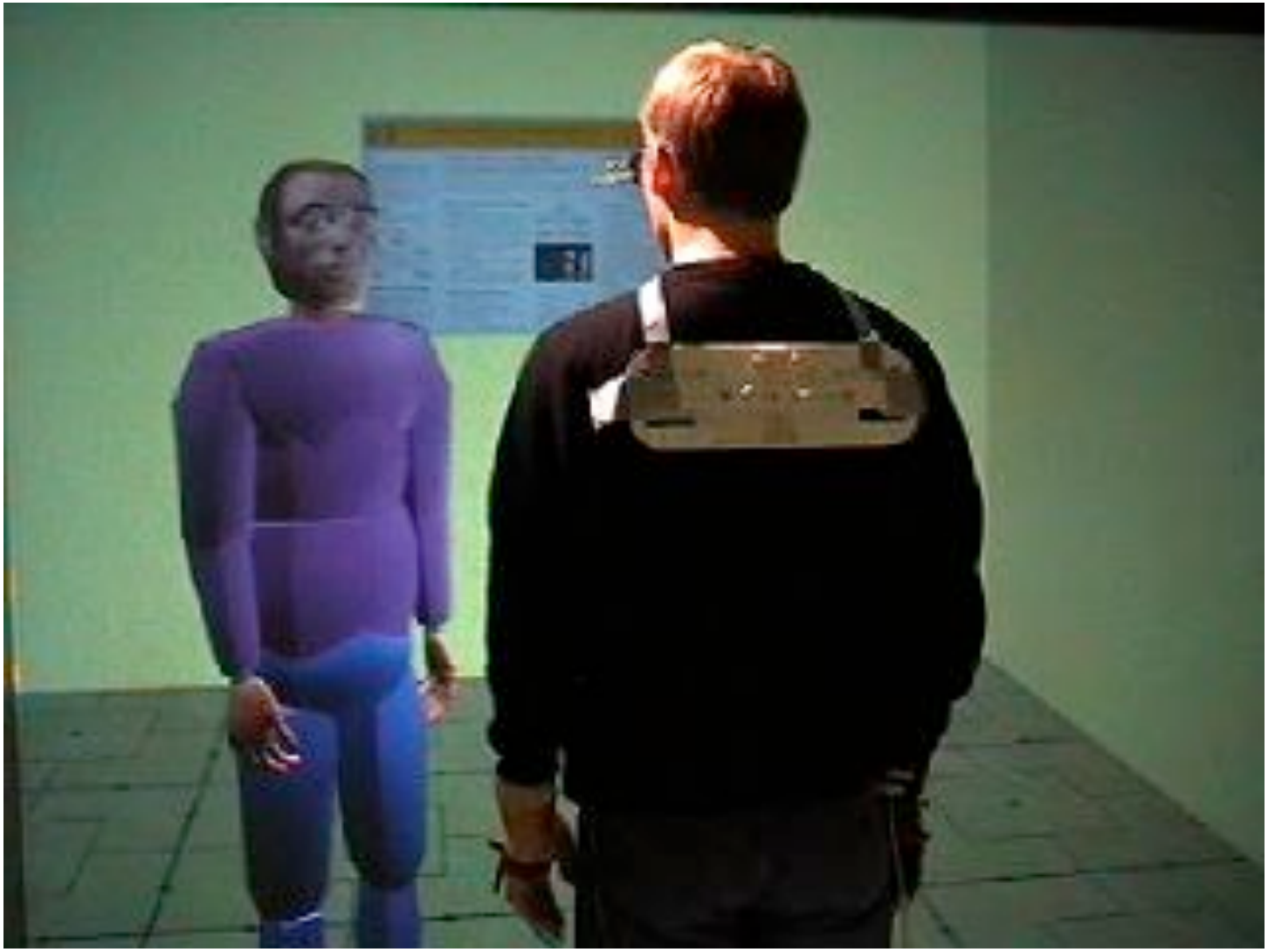
Imitation von Gestik

Mensch macht vor, MAX macht nach

Analyse der Gesten mit HamNoSys

HamNoSys als Ausgangspunkt für das Nachmachen

Echtzeit!



Max als Museumsführer

→ seit Januar 2004 arbeitet Max im Heinz Nixdorf MuseumsForum, Paderborn

← "Wie ist das Wetter in Hamm?"

→ "Max, das bin ich. Ich bin eine künstliche Person, die sprechen und gestikulieren kann. Ich selber bin künstlich, kann mich aber genau so ausdrücken wie du..."

Promis und Max

NR. 154, MITTWOCH, 6. JULI 2005

Kennst du Talle? *Ipke Wachsmuth von der Uni Bielefeld (r.) half Kanzler Gerhard Schröder, Avatar Max zu befragen.* FOTO: TOBIAS HEYER

Max auf MTV

Max – ein autonomer Agent

(vergleiche Wooldridge & Jennings)

- **Autonomie:** agiert ohne direkten Eingriff von außen, hat eigene Kontrolle
- **Soziale Fähigkeiten:** interagiert mit dem Menschen (sprachlich und nichtverbal)
- **Reaktivität:** nimmt seine Umgebung wahr und kann darauf reagieren
- **Proaktivität:** hat zielorientiertes Verhalten und kann Initiative ergreifen

Was ist Künstliche Intelligenz?

Künstliche Intelligenz ist die Untersuchung von geistigen Fähigkeiten durch Berechnungs-modelle.

PERCEIVE
REASON
ACT

Künstliche Intelligenz ist die Untersuchung von Berech-nungsverfahren, die es ermöglichen, wahrzunehmen, zu denken und zu handeln.

Künstliche Intelligenz befasst sich mit der Konstruktion informa-tionsverarbeitender Systeme, die kognitive Leistungen modellieren und in technischen Anwendungen verwerten.

Einige Abschlussbemerkungen

Der Intelligenzbegriff der KI ist "computational" (Verarbeitung von Information durch Symbole).

KI-Theorien sind "generativ" in dem Sinne, daß sie Synthese vor Analyse stellen.

Paradigmen ändern sich:

- **eingebettete Systeme (die mit Sensoren und Effektoren in ihre Umgebung gekoppelt sind): Situierte KI**
- **anstelle "globaler" Rekonstruktion intelligenten Verhaltens Betrachtung simpler interagierender Systeme: Agenten-Systeme, Verteilte KI**

Kann KI eine Weltformel für Intelligenz liefern?

Kein einzelner Ansatz kann alle Merkmale von Intelligenz reproduzieren / erklären.

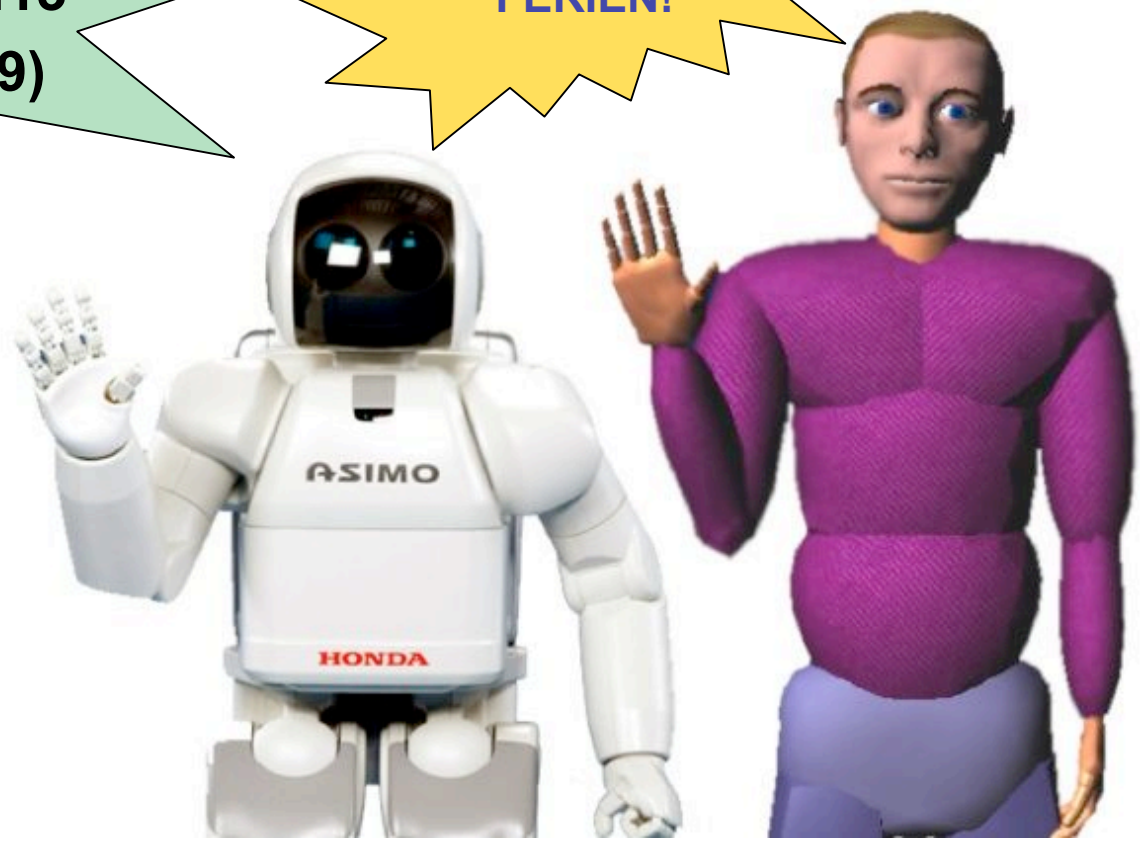
Minsky (in einer 93er Ausgabe von Scientific American):

"The mind is a tractor-trailor, rolling on many wheels, but AI workers keep designing unicycles."

Danke für Ihre Aufmerksamkeit!

Viel Erfolg bei der Klausur:
17.02.2009, 10-12 Uhr in H13
(Anmeldung bis 06.02.2009)

SCHÖNE FERIEN!



– Ende –