Universität Bielefeld

9 Ausgewählte Themen

29. Vorlesung: Virtuelles Konstruieren mit Gestik und Sprache

(Folien in Auswahl)

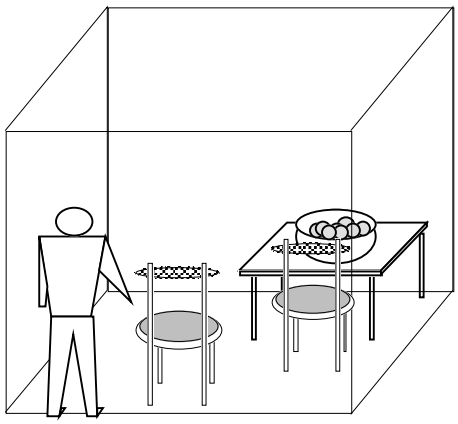
Methoden der Künstlichen Intelligenz

Ipke Wachsmuth WS 1999/2000

Universität Bielefeld

Anthropomorphe Interfaces

- Schnittstellen nach dem Bild des Menschen
- natürliche Mensch-Maschine-Kommunikation
- hochaktuell im Bereich interaktiver Multimedia
- Gestik, Sprache, sprachbegleitende Gestik



Stell diesen Stuhl da hin!

Kommunikation in der Situation

Zentral in unseren Arbeiten:

- Multimedia
- Multimodalität
- Multiple Repräsentationen
- Multi-Agenten-Systeme

Virtuelles Konstruieren

- Innovatives Verfahren zur Entwurfsunterstützung und Montage-Erprobung in Frühphasen industrieller Herstellungsprozesse
- Erprobung von Entwurfsvarianten und Fertigungs-plänen auf Basis von CAD–Modellen mechanischer Objekte in der virtuellen Realität
- Überprüfung konstruktiver und gestalterischer Merkmale am virtuellen Prototypen eines geplanten Produkts (Kostenersparnis, Qualitätssicherung)

Probleme in der 3D-Grafik

- mühsame Handhabung synthetischer Modelle
- mangelnde Unterstützung des interaktiven Entwurfs
- Fehlen menschengerechter Kommunikationsformen

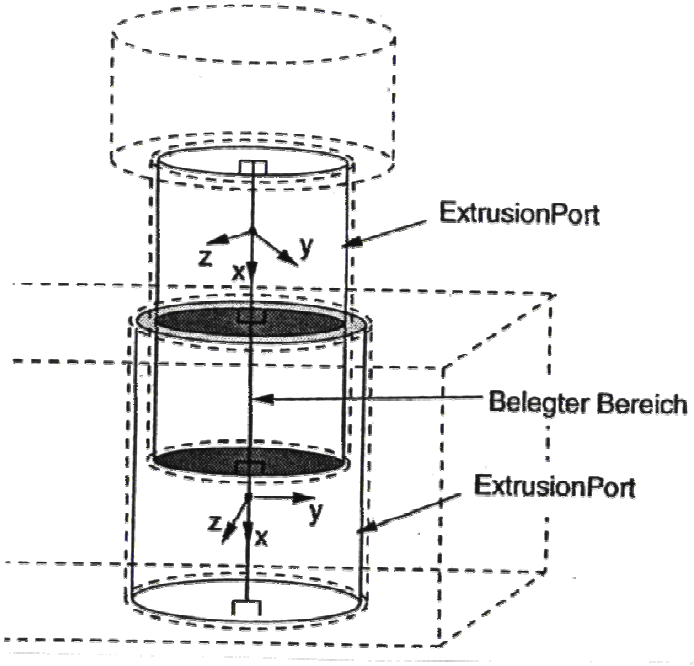
Wünsche:

- physikgerechte "reale" Manipulationsmöglichkeiten
- interaktive Systeme, die Benutzereingriffe in einer visualierten Szene unmittelbar umsetzen

Virtueller Konstrukteur

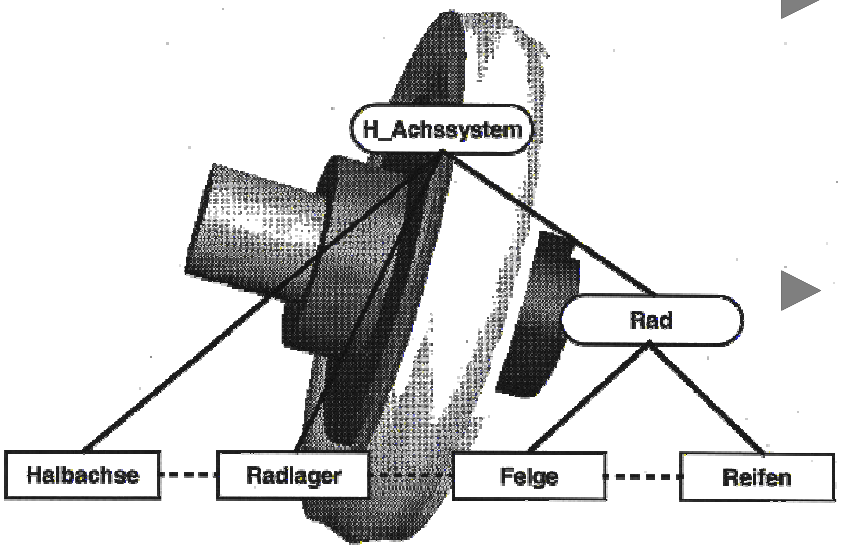
entwickelt im Projekt CODY - SFB 360

- Wirklichkeitsgetreues Fügen: bewegen, verbinden, trennen
- Vereinfachung durch das Einschränken von Freiheitsgraden
- physikalisch korrekte Relativbewegungen der Aggregate

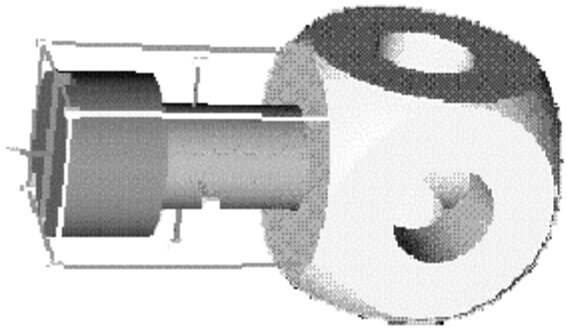


Dynamische Konzeptualisierung

Arbeiten mit Bernhard Jung und Martin Hoffhenke, Projekt CODY - SFB 360



- Schritthaltender Aufbau begrifflicher Objektrepräsentationen vor dem Hintergrund einer Montagewissensbasis
- Abmessen der metrischen Portkonsumptionen direkt im Geometriemodell



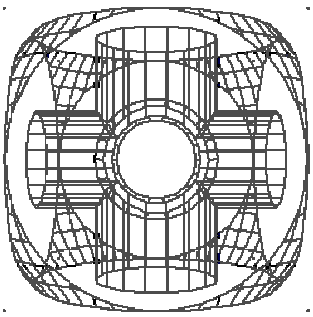
Erweiterte Freiheitsmatrix Schraube-1:

$$\begin{pmatrix} 3.3696 & 18.6304 & \text{FREE} & \text{FREE} \\ \text{BLOCKED} & \text{BLOCKED} & \text{BLOCKED} & \text{BLOCKED} \\ \text{BLOCKED} & \text{BLOCKED} & \text{BLOCKED} & \text{BLOCKED} \end{pmatrix}$$

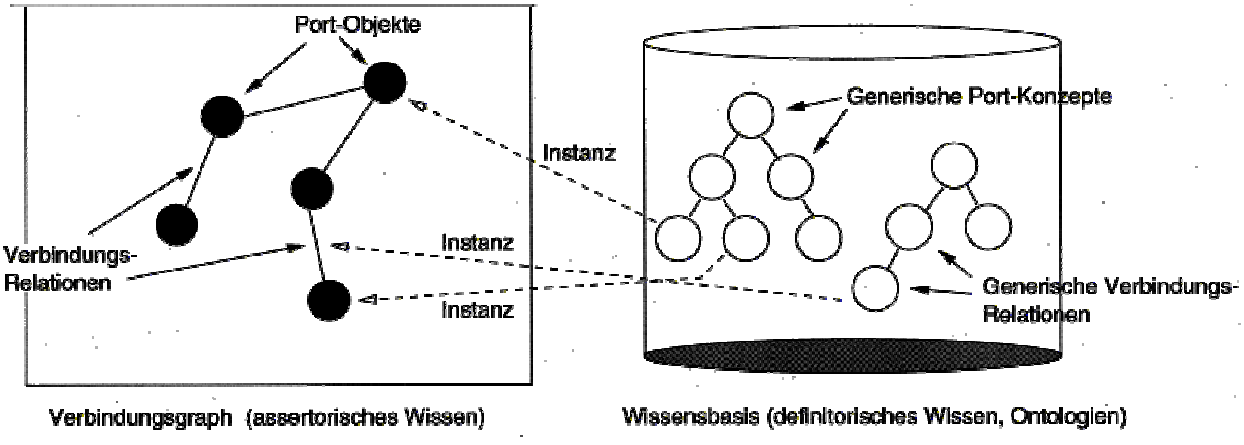
{(0,0,2,0, 0.8),(1,0,3,0, 0.8)}

Virtueller Konstrukteur

Arbeiten mit Stefan Kopp im Projekt CODY - SFB 360

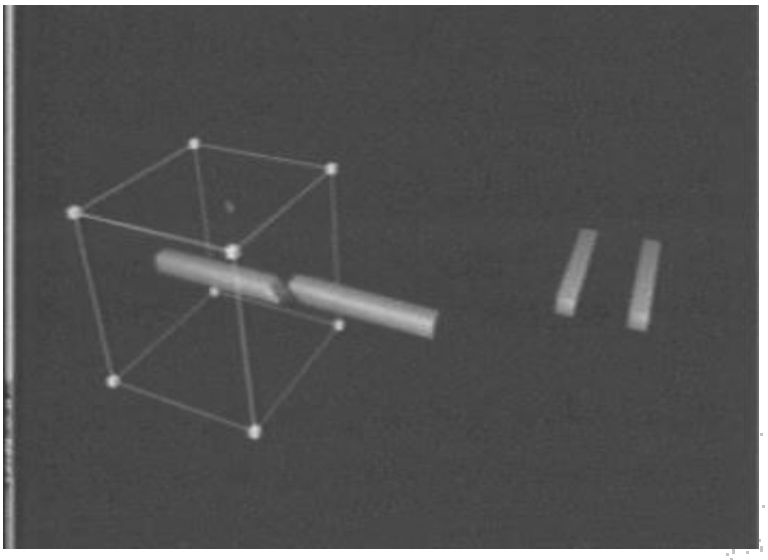


- Simulation korrekten Objektverhaltens durch dynamisch anpaßbare Strukturbeschreibungen



Verbindungen durch “Ports”

gezeigt: Plane-Ports

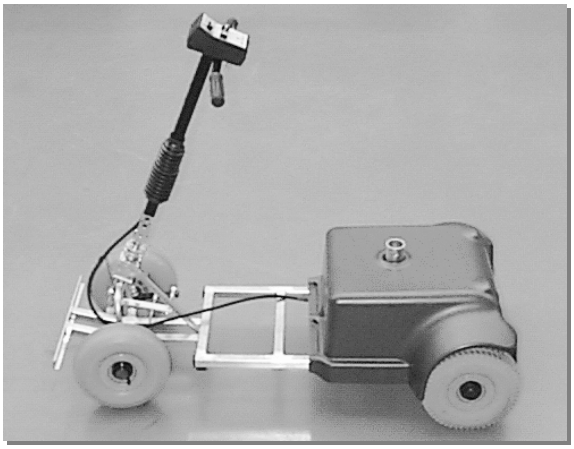


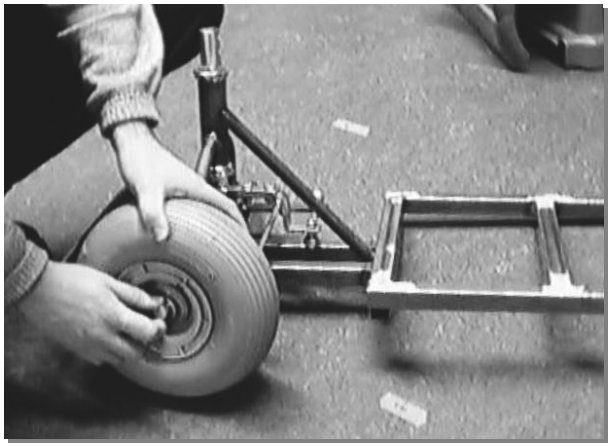
Plane-Ports, Point-Ports

Extrusion-Ports

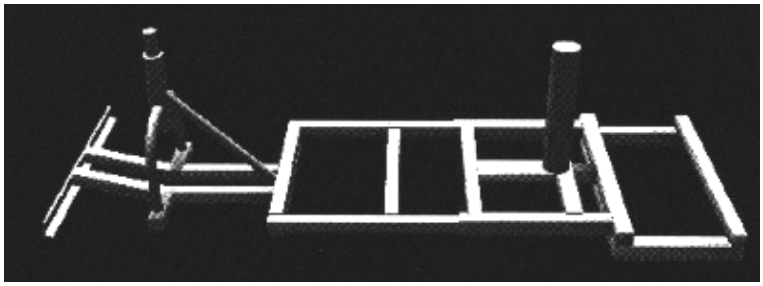
- Kollisionserkennung
- Zusammenschnappen
- Freiheitsgrade einschränken
- Verbindungen fixieren (Material- oder Kraftschluß)

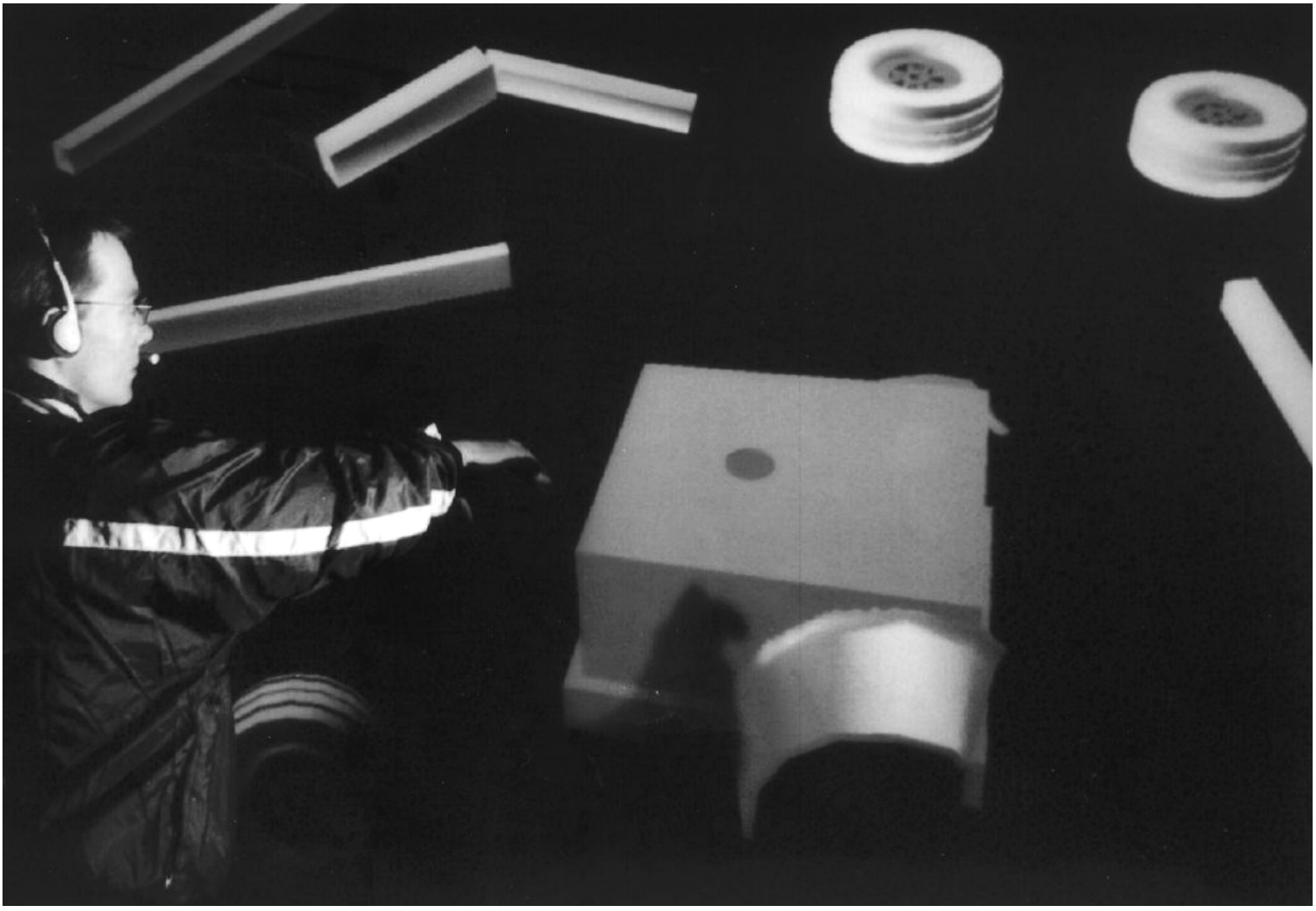
Konstruktionsziel: City-Mobil





Rahmen (virtuell)





Universität Bielefeld

Gestenerkennung

Arbeiten mit Marc Latoschik, Projekt SGIM

turn
(mimetisch)

select
(deiktisch)

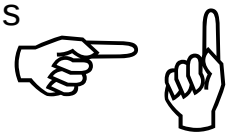
drag —— and —— drop

"Gesten"

Körperbewegungen, die bedeutsame Information übermitteln

Vier Typen (zunächst) interessant für MMK:

1. **emblematische Gesten (Symbole):** Bedeutung per Konvention festgelegt
2. **deiktische Gesten:** Handlungen des Zeigens auf ein Objekt oder eine Region



3. **mimetische Gesten:** Handlungen, die die Interaktion mit einem Objekt imitieren
4. **koverbale Gesten:** Gesten, welche das Sprechen begleiten

➡ Typ1-Gesten stark erforscht in der MMK; wir fokussieren auf die Typen 2 bis 4; zusätzlich: Blickrichtung, basiert auf der Kopforientierung.

Ziel: Multimodale Interaktion

Arbeiten mit Marc Latoschik, Projekt SGIM

SGIM Speech and Gesture Interfaces for Multimedia

Virtuelle Wissensfabrik

NRW

Gestenerkennung – Ansätze

Ansatz 1: Basisinteraktionen, direktes Feedback

- ★ Zeigen
- ★ Greifen
- ★ Loslassen
- ★ Translation
- ★ Rotation

Ansatz 2: Formbasiertes Vorgehen

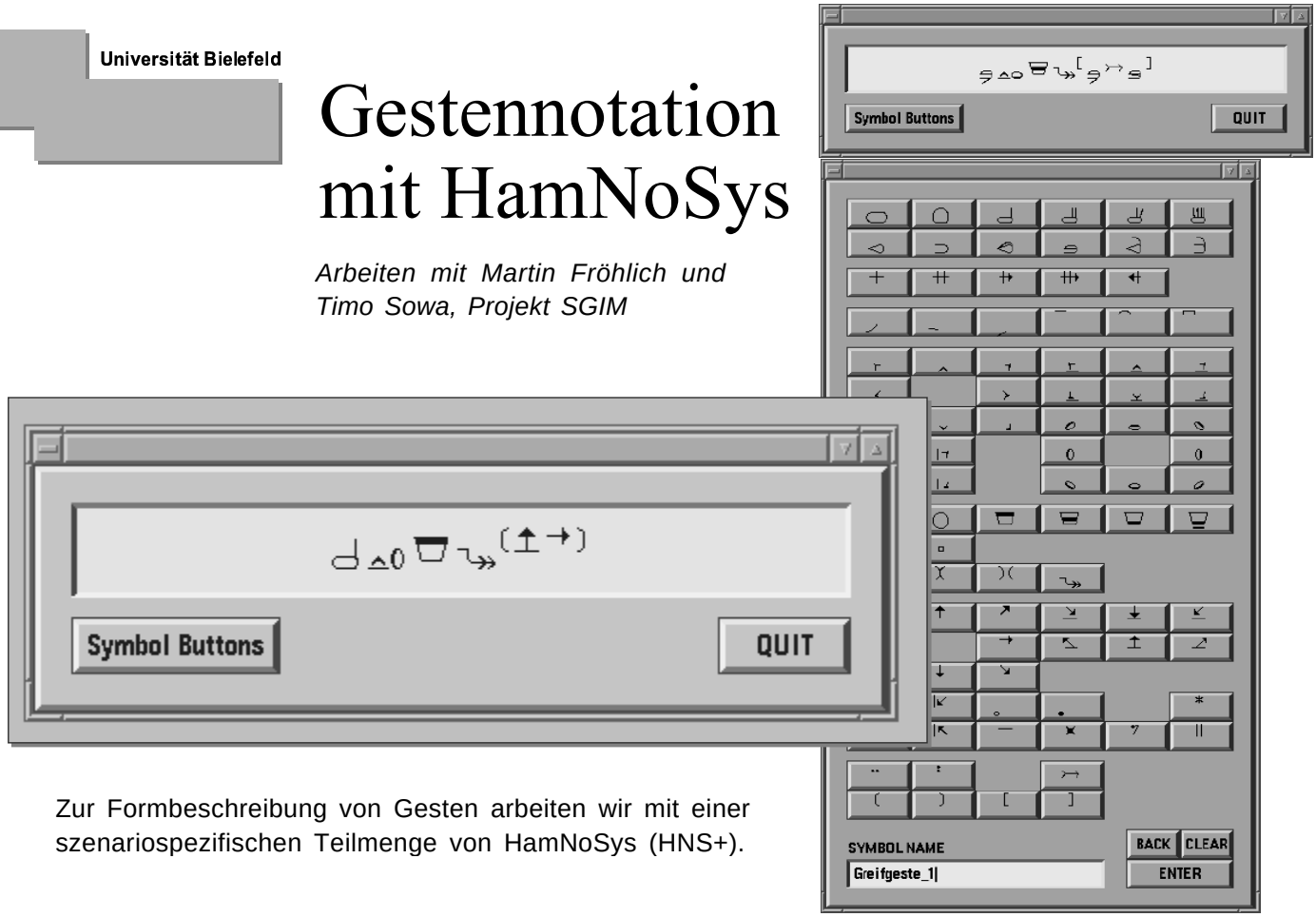
- ◆ expliziter (wissensbasierter) Ansatz
- ◆ atomare Formelemente von Gesten werden zu 'Gesten-Worten' zusammengesetzt

Symbole für Körperteile

Universität Bielefeld

Gestennotation mit HamNoSys

Arbeiten mit Martin Fröhlich und Timo Sowa, Projekt SGIM



Zur Formbeschreibung von Gesten arbeiten wir mit einer szenariospezifischen Teilmenge von HamNoSys (HNS+).

Universität Bielefeld

```
(defrule rule32
?tmp28 <- (object (is-a MoveA))    ;    eine Handvorwärtsbewegung
?tmp31 <- (object (is-a MoveR))    ;    eine Handrechtsbewegung
=>
(SEQUENCE tmpcl25 ?tmp28 ?tmp31)) ;    in Sequenz
```

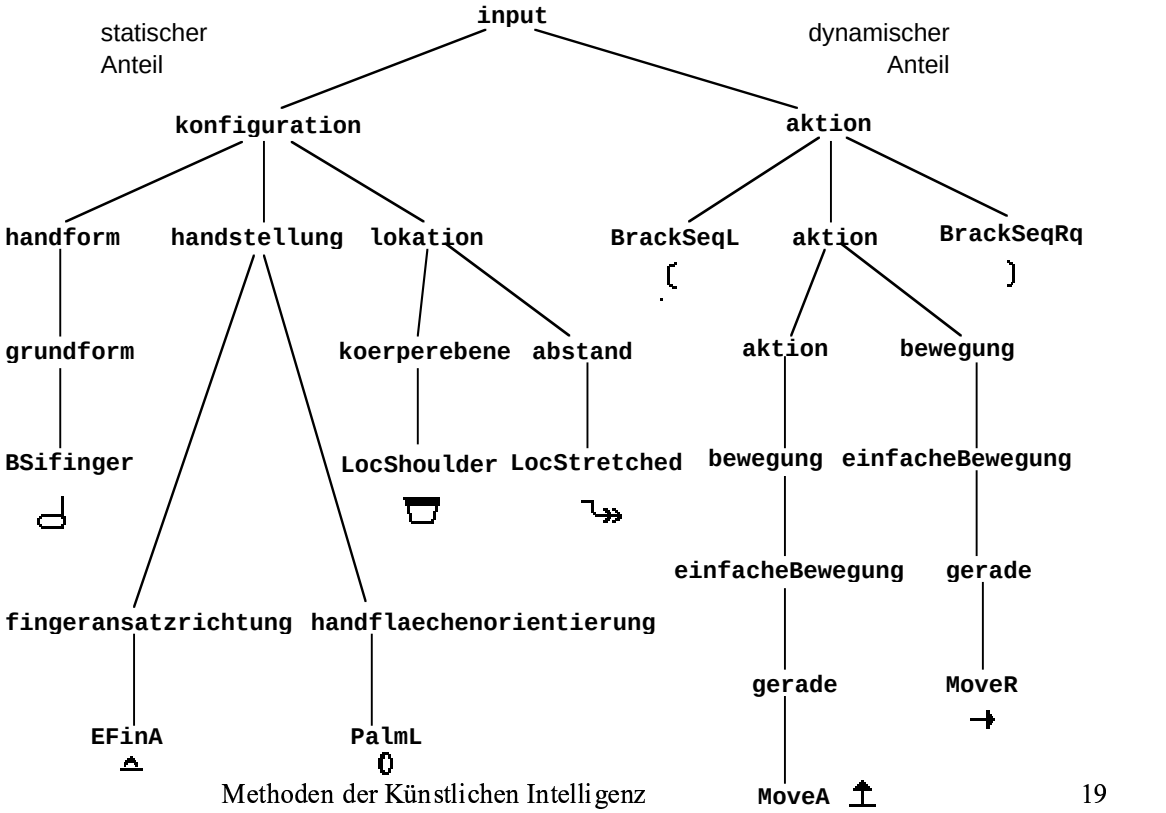
Hier ist automatisch eine CLIPS-Regel `rule32` generiert worden, die ein Token der Klasse `tmpcl25` assertieren kann, das eine erkannte Sequenz von bestimmten Handbewegungen registriert.

```
(defrule rule36
?tmp16 <- (object (is-a BSifinger))    ;    Zeigefinger gestreckt
?tmp19 <- (object (is-a EFinA))        ;    nach vorn gestreckt
?tmp22 <- (object (is-a PalmL))        ;    Handfläche links
?tmp23 <- (object (is-a LocShoulder)) ;    in Schulterhöhe
?tmp24 <- (object (is-a LocStretched))     ;    ausgestreckt
?tmp35 <- (object (is-a tmpcl 25))         ;    bewegt wie oben
=>
(PARALLEL DUMMY ?tmp16 ?tmp19 ?tmp22 ?tmp23 ?tmp24 ?tmp35))
```

Hier ist automatisch eine Regel `rule36` generiert worden, die ein Token der Klasse `DUMMY` assertieren kann, das das Auftreten einer bestimmten Handpostur in Verbindung mit oben beschriebener Bewegung registriert.

(für `DUMMY` kann ein beschreibender Gestenname gesetzt werden)

HNS+ Parse Tree für



Multimodale Integration

Zwei "logistische" Probleme zu lösen (Srihari, 1995):

- **Segmentierungsproblem**

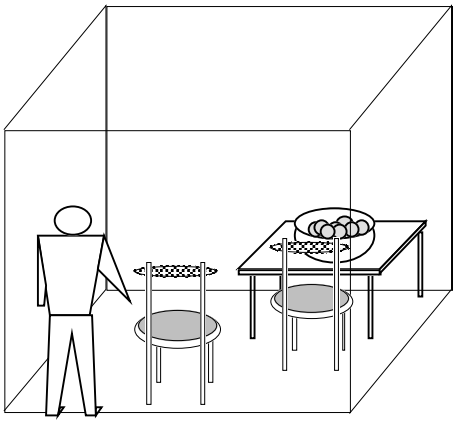
  Wie sind die Prozeßeinheiten zu determinieren, die das System in einem Zyklus verarbeiten soll?

- **Korrespondenzproblem**

  Wie sind die Querbezüge zwischen multiplen Modalitäten (Sprache/Gestik) zu determinieren?

Korrespondenzproblem

- Wie sind die Querbezüge zwischen Sprache und Gestik zu determinieren?



Stell diesen Stuhl da hin!

Wichtigkeit der Rekonstruktion zeitlicher Korrespondenz – Beispiel:

- **Stell** <Zeigegeste> **<u>diesen</u> Stuhl da hin !**
  Spezifikation des Referenten
  (Presupposition: Zielort bekannt)
- **Stell diesen Stuhl** <Zeigegeste> **<u>da</u> hin !**
  Spezifikation des Zielorts
  (Presupposition: Referent bekannt)

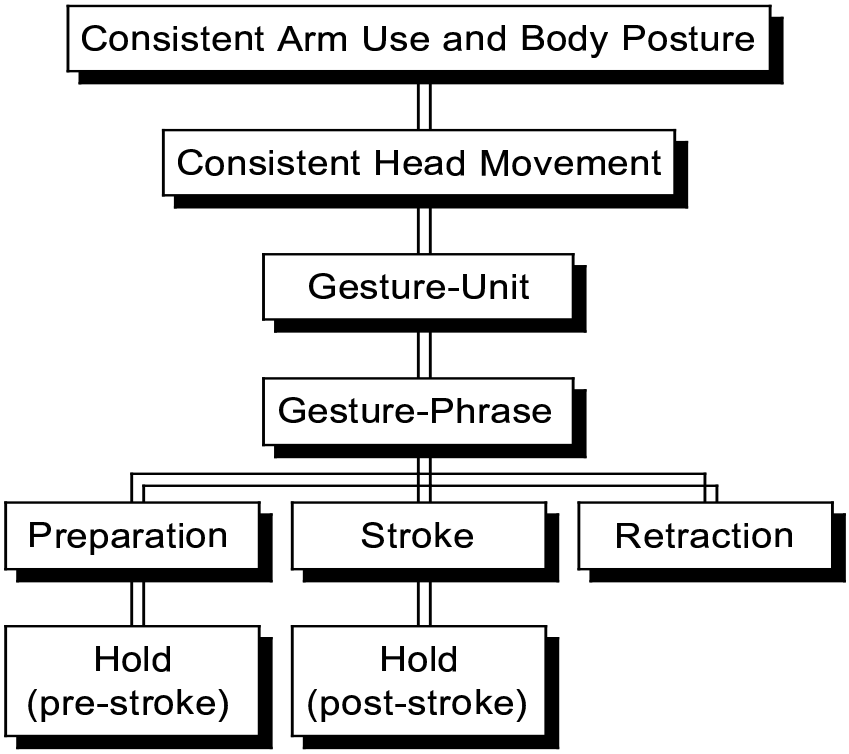
Gestik und Sprache

Gestik und Sprache sind eng gekoppelt

Rhythmische Korrelation von Sprechtakt und Gesten-Stroke

Gesten: Kinetische Struktur

Nach McNeill, Levy & Pedelty (1990)



Zeitliche Kopplung

- Der sog. Gesten-Stroke ist oft durch einen abrupten Halt gekennzeichnet, der mit den gesprochenen Wörtern synchronisiert ist
- der Stroke tritt nicht nach dem betonten Wort auf, sondern gleichzeitig oder kurz vorher

=> Hypothesen zur rhythmischen Signalregistrierung für Symbolextraktion und multimodale Integration

Sprechtakt: “stress timing”

(in germanischen Sprachen)

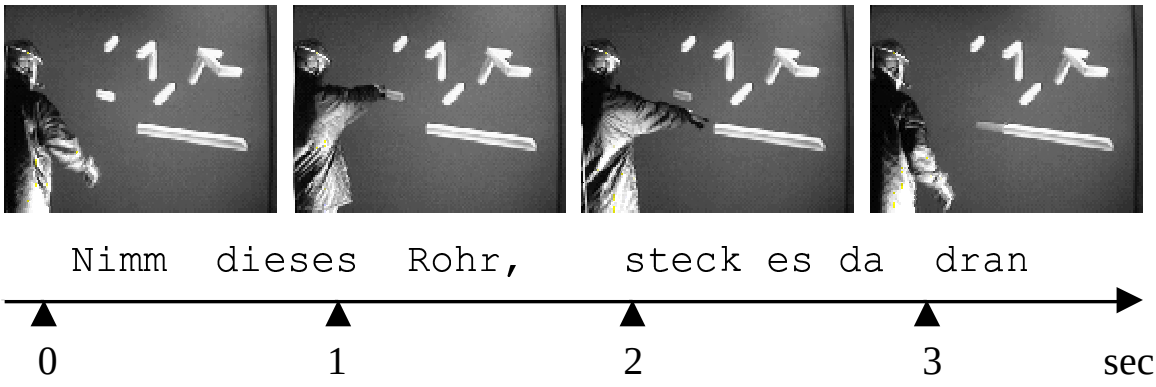
- Betonung ("stress") durch zeitliche Dehnung einzelner – und Verdichtung anderer – Silben:

Der	Zug	nach	Köln
Die	Zü-ge	nach Ber-	lin
Die	Ei-sen-bahn	nach	Eisenach
Der Inter-	Ci-ty	nach Bad	Segeberg

=> relativ konstante Dauer zwischen betonten Silben, unabhängig von Silbenanzahl

(beim gleichen Sprecher unter gleichen Bedingungen)

Temporale Integration



aus der laufenden Arbeit...

Literatur

- Ein Überblicksartikel zum Thema „Virtuelles Konstruieren mit Gestik und Sprache“ kann bei Interesse zur Verfügung gestellt werden.