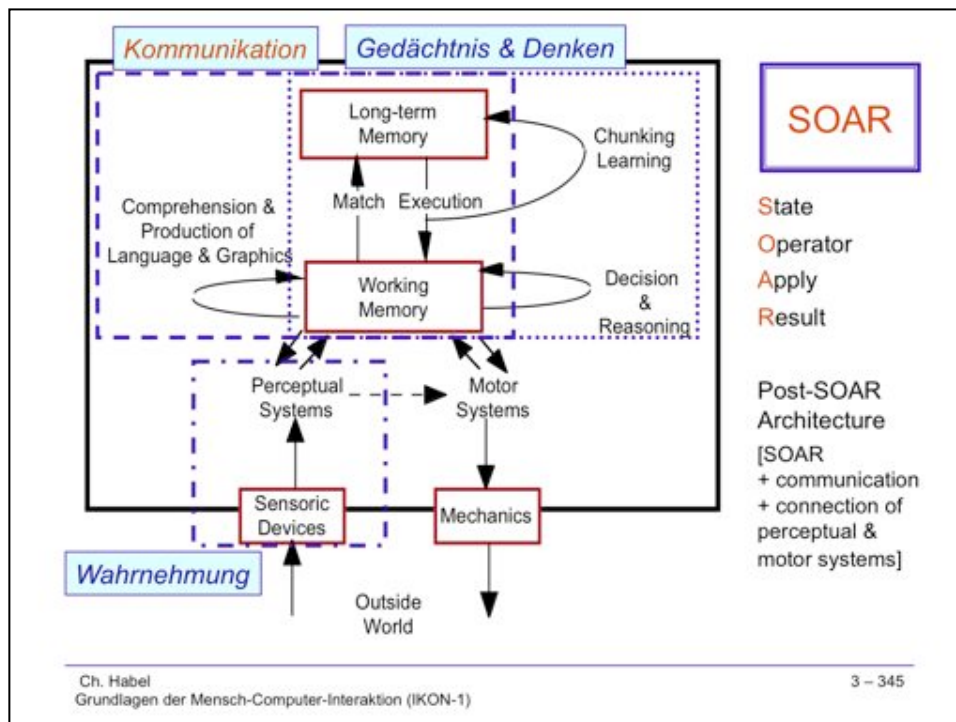


Grundlagen der Mensch-Computer Interaktion

3. Kapitel

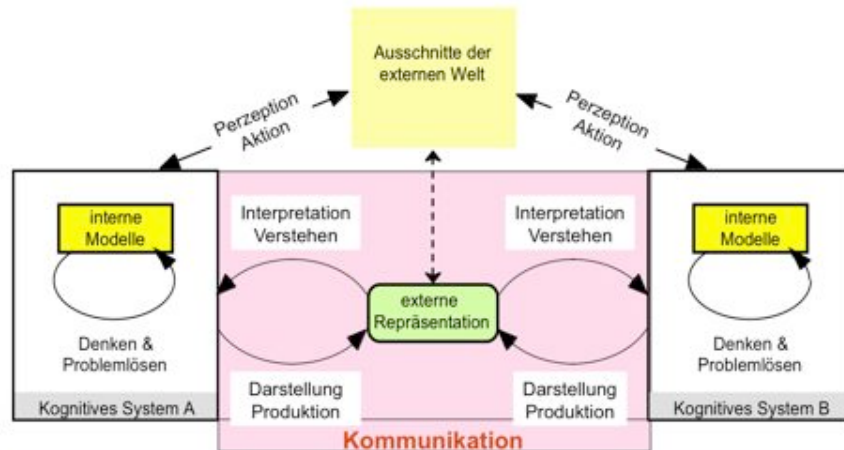
Der Mensch

- Ein- und Ausgabe: Wahrnehmung und Handeln
 - Wahrnehmung: Modalitäten der Wahrnehmung
- Gedächtnis, Denken & Problemlösen
- Kommunikation & Interaktion
 - Prinzipien der Kommunikation
 - Sprachliche Kommunikation
 - Multimodale Kommunikation



- entspricht weitgehend Folie 3-3.
- Im abschliessenden Teil es 3. Kapitel befassen wir uns mit Kommunikation, d.h. Vermittlung von Information / Wissen durch externe Repräsentationen

Kommunizierende und interagierende Agenten



Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 – 346

- Im Rahmen des Informationsverarbeitungsparadigmas werden *Kognitive Systeme* als Systeme aufgefasst, deren erfolgreiches Handeln in ihrer Umwelt (=externe Welt) wesentlich darauf beruht, dass sie über interne Modelle (= interne Repräsentationen) verfügen. Denken und Problemlösen kann als Verarbeitung interner Modelle angesehen werden (-> vorangehenden Abschnitt dieses Kapitels).
- Über interne Repräsentationen hinaus, sind gerade für den Menschen, externe Repräsentationen von grosser Bedeutung, z.B. Sprache, Zahlssysteme, Diagramme,, die einerseits beim Problemlösen und Denken und andererseits in der Kommunikation mit anderen Menschen (oder auch mit Computern) das „Interface“ zwischen den Kommunikationspartnern darstellen.
- Erfolgreiche Kooperation zwischen Agenten setzt voraus, dass das Handeln der beteiligten Agenten auf einander abgestimmt wird. Dies kann dadurch erreicht werden, dass die Agenten miteinander kommunizieren, erfordert aber auch, dass die Handlungen der anderen beobachtet werden, und sogar Prognosen über zukünftige Handlungen aufgestellt und berücksichtigt werden.

Kommunikation

- Funktion von Kommunikation
 - Übertragung von Erfahrung und Wissen
 - Koordination: Basis kooperativen Handelns
- Modalitäten der Kommunikation
 - Sprache
 - primär menschliche Fähigkeit
 - (evolutionär) erfolgreichste Kommunikationsform
 - ist das leistungsfähigste (generelle) Repräsentationssystem
 - Bildhafte und graphische Kommunikation
 - Landkarten, Skizzen, Zeichnungen
 - Graphen, Diagramme, ...
 - Gesten, Mimik, taktile Kommunikation
 - Kommunikation ist häufig **multimodal**

Modalitäten der Kommunikation korrespondieren zu “repräsentationellen Modalitäten”. Dies ist NICHT identisch mit sensorischen Modalitäten (Modalitäten der Wahrnehmung), die zu Beginn des Kapitels 3 besprochen wurden.

Kommunikation und Kognition: Informatik-Perspektiven

- Mensch-Computer Interaktion & Schnittstellengestaltung
 - natürlich-sprachliche Schnittstellen
 - graphische Schnittstellen
 - Masken
- Maschinelle Verfahren des Verstehens und Produzierens von externen Repräsentationen
 - Maschinelle Sprachverarbeitung
u.a.: Übersetzung, Informationsextraktion,...
 - Bildverstehen & Visualisierung, einschl. Virtuelle Realität
- Designprinzipien für die Entwicklung von maschinellen Kommunikationssystemen
 - Kommunikation zwischen autonomen Agenten
- Assistenzsysteme für Menschen mit Behinderungen / Beeinträchtigungen (des Sehens, des Hörens, des Sprechens, des Zugriffs auf das mentale Lexikon,...)

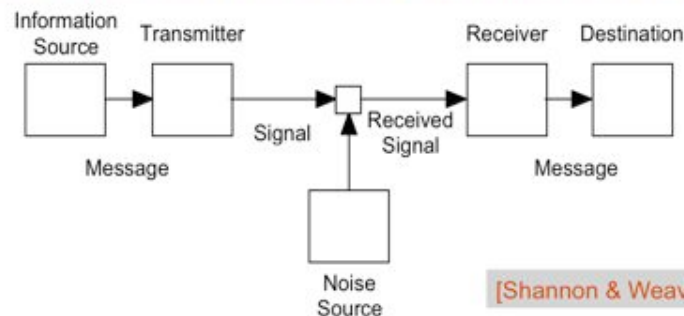
Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 – 348

Behinderungen

- des Sehens (insbesondere Blinde): Text-to-speech-Systeme -> Vorlesen für Personen mit Sehstörungen
- des Hörens (Gehörlose, bzw. Schwerhörige): Unterstützung für den Erwerb und den Gebrauch von Gestensprachen (ASL, DGS, ...)
- des Sprechens („Taubstumme“: s.o.): text-to-speech Assistenz
- des Zugriffs auf das mentale Lexikon und die Grammatik: Rehabilitation & Training von Aphasikern.

Eine zweite Sichtweise auf Kommunikation: *The mathematical theory of communication*



[Shannon & Weaver, 1949]

- Level A: How accurately can the symbols of communication be transmitted? (The technical problem) **Data**
- Level B: How precisely do the transmitted symbols convey the desired meaning? (The semantic problem) **Information Knowledge**

Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 – 349

- Die mathematische Informationstheorie / mathematische Kommunikationstheorie
 - ist zentral für die Signalverarbeitung und die Technische Kommunikation, z.B. in den Bereichen Telekommunikation, Datenübertragung per Netz, Funk, Satellit,...
 - fokussiert auf andere Fragestellungen und legt dabei eine andere Konzeption von Kommunikation zugrunde (vgl. die folgende Folie)
 - wird in der Vorlesung Rechnerstrukturen behandelt.
- Shannon, C. E. (1948). A mathematical theory of communication. Bell Sys. Tech. Journal 27: 379–423, 623–656.
- Shannon, Claude E. & Weaver, Warren (1949). The mathematical theory of communication. Urbana, IL: University of Illinois Press.

Shannon's *Mathematical Theory of Communication*

- Shannon (1949, p. 31):

"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. **Frequently** the **messages have meaning**; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These **semantic aspects of communication are irrelevant to the engineering problem**. The significant aspect is that the actual message is one *selected from a set* of possible messages...."

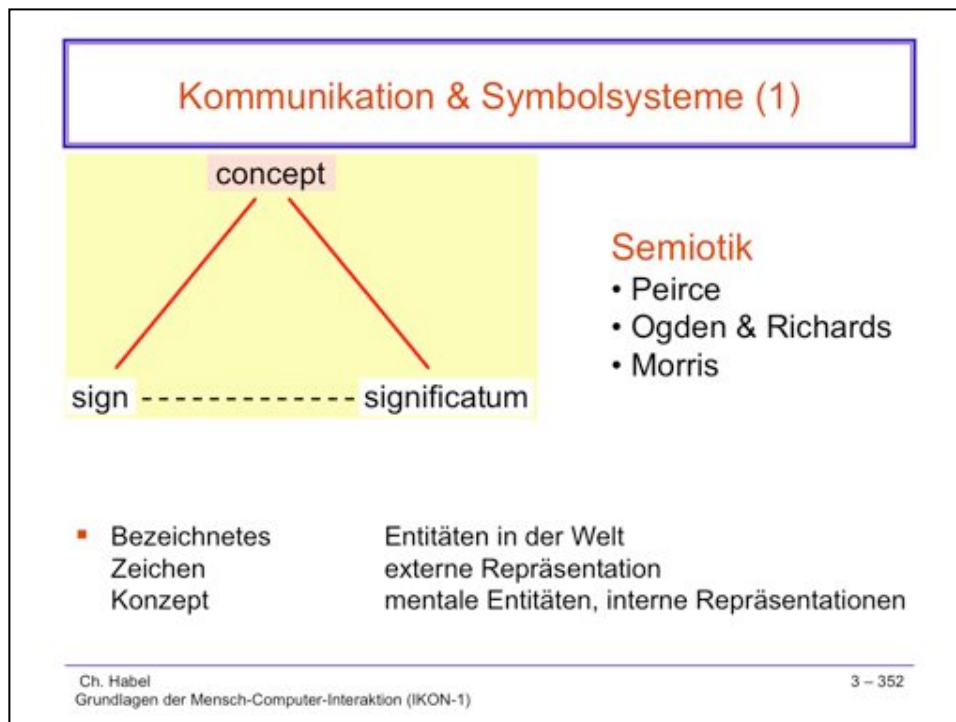
- **Data level** vs. **information / knowledge level**

- Wie die Originalarbeit von Shannon zeigt, ist die Konzeption von *Bedeutung* nicht im Fokus der Mathematischen Kommunikationstheorie.
- Im Gegensatz hierzu, ist für den Bereich der Mensch-Computer-Interaktion die *Bedeutung* von Äusserungen, Graphiken (und anderen Repräsentationen), sowie die Bedeutung von Handlungen (z.B. Gesten) der zentrale Aspekt bei der Untersuchung von externen Repräsentationen und Interaktionshandlungen.

(Sprachliche) Kommunikation und Bedeutung: Die zentrale Fragen

- Was ist die Bedeutung von sprachlichen Äusserungen?
 - Formale Repräsentationen der Bedeutung natürlich-sprachlicher Äusserungen (Formale Semantik)
- Wie werden Bedeutungen kodiert?
 - Die – gegebenenfalls situationsabhängige – Berechnung von Äusserungen aus Bedeutungen: Sprachproduktion.
- Wie werden – gegebenenfalls situationsabhängig – Bedeutungen aus sprachlichen Äusserungen berechnet?
 - Sprachverstehen
- Die entsprechenden Fragen stellen sich für alle Modi der Kommunikation, z.B. für
 - graphische Kommunikation
 - Kommunikation über Masken

- Sprache hat eine zentrale Position in der wissenschaftlichen Untersuchung von Kommunikation (ähnlich der Stellung der visuellen Wahrnehmung für die Perzeptionsforschung)
- In diesem Sinne sind die drei – oben aufgeführten – Fragen zu Bedeutung und sprachlichen Äusserungen als Spezialfälle zu Bedeutungen und Kommunikation (im Allgemeinen) zu sehen.



Konzeption der scholastischen Philosophie (spätes Mittelalter)
 „vox significat [rem] mediantibus conceptibus“
 „Das Wort bezeichnet [das Ding] über vermittelnde Konzepte.“

Semiotik – Allgemeine Theorie der Zeichen

Peirce, C. S. (1931–58). *Collected Papers*. Cambridge, MA: Harvard University Press.

- Peirce' Terminologie wird heute viel verwendet, jedoch häufig nicht entsprechend zu seinen Charakterisierungen.

Saussure, F. de (1916). *Cours de linguistique générale*. Lausanne-Paris: Payot.

Ogden, C.K. & Richards, I.A. (1923). *The Meaning of Meaning*. Routledge & Kegan Paul: London.

Morris, C. (1938). *Foundations of the Theory of Signs*. University of Chicago Press: Chicago.

Eco, U. (1976). *A Theory of Semiotics*. Bloomington, IN: Indiana University Press.

Von Eckhardt, Barbara (1993). *What is Cognitive Science*. Cambridge, MA: MIT Press.

Chandler, Daniel (1994): *Semiotics for Beginners*

[WWW document] **URL** <http://www.aber.ac.uk/media/Documents/S4B>

de Souza, Clarisse Sieckenius (2005). *The semiotic engineering of human-computer interaction*. Cambridge, MA: MIT Press.

- de Souza hat eine komplette Konzeption der Mensch-Computer-Interaktion auf dem Fundament der Semiotik aufgebaut.

Icons, indexes and symbols

- Wörter und Sätze sind Symbole, z.B. das Verb „öffnen“

- Piktogramme, z.B. Verkehrszeichen, können ikonisch oder symbolisch sein, oder Kombinationen von beidem



- EARCONs
 - sind Indizes nicht Icons



open

Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 – 353

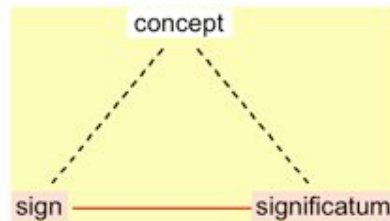
- Aus der Sicht der Peirce 'schen Semiotik sind *Symbole* ein spezieller Typ von *Zeichen*, der im Bereich der Kommunikation mit Sprache im Zentrum steht.
 - Ikons* weisen eine (visuell bildhafte) Ähnlichkeit zu den Objekten, die sie repräsentieren auf.
 - Indizes* stellen verweisen im durch den situativen Kontext vermittelt auf das Objekt, das sie repräsentieren.
- EARCONs sind meist keine akustischen "Gegenstücke" zu ICONs, obwohl die Wortschöpfung am Begriff Icon angelehnt ist. Es handelt sich überwiegend um Indices, die verwendet werden, um durch Töne / Klänge / Tonfolgen Bedeutung zu vermitteln.

Einen ausführlichen Einblick in Forschung zu und Anwendung von Earcons erhalten Sie auf der Web-Site von Stephen Brewster's *Multimodal Interaction Group*: <http://www.dcs.gla.ac.uk/~stephen/>

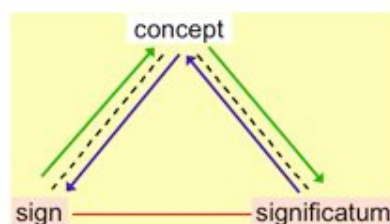


- Ein ikonisch-symbolisches Piktogramm

Kommunikation & Symbolsysteme (2)



- **Referenzbeziehung:**
Zeichen – Bezeichnetes
 - willkürlich, konventionalisiert
 - *Hund – dog*



- Zeichen in einer – erfolgreichen – Kommunikation
 - Der **Sender** produziert ein Zeichen
 - Der **Empfänger** interpretiert das Zeichen (induziert ein Konzept mit Beziehung zum Bezeichneten)

Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 – 355

- Die Beziehung (Referenz) zwischen *Zeichen (Symbol)* und *Bezeichnetem*, d.h. dem Objekt in der Welt, wird in einer Kommunikationssituation durch die Teilnehmer, etwa die Sprecher und Hörer hergestellt.
- Im Fall einer sprachlichen Kommunikation beabsichtigt z.B. die Sprecherin über ein Objekt (significatum) zu sprechen; sie verwendet ihr semantisches Gedächtnis (genauer ihr konzeptuellen Repräsentationen) um das Objekt einer Kategorie zuzuordnen und hat über das Konzept den Zugriff zu einem Wort ihrer Sprache, das sie äussert (spricht), und somit eine externe Repräsentation (das Zeichen) realisiert. [die ist in der unteren Abb. durch die beiden schwarzen Pfeile symbolisiert.
Der Hörer hört das Wort, aktiviert das entsprechende Konzept, und versucht eine Instanz der zugehörigen Kategorie in der Realität zu finden, auf die die sprachliche Beschreibung „passt“.
Die Kommunikation ist erfolgreich, wenn der Hörer das von der Sprecherin „gemeinte“ Objekt indentifiziert.

Mensch-Roboter Interaktion

ein Spezialfall der Mensch-Computer Interaktion

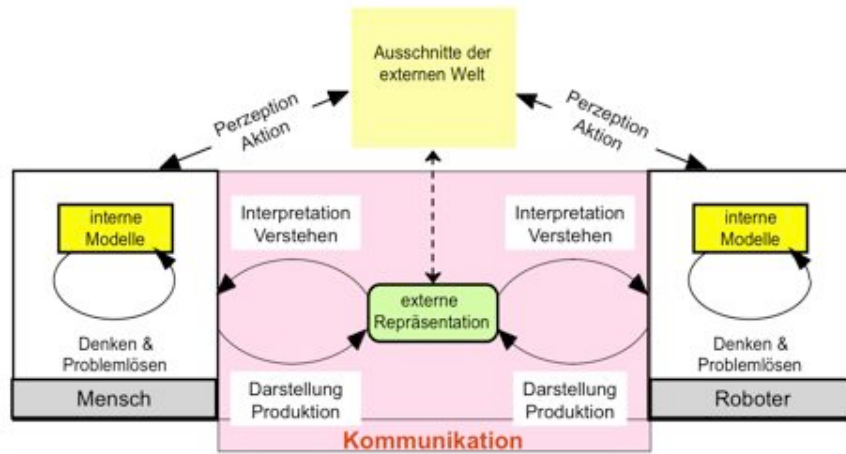
Unterschiede zur Mensch-Computer Interaktion

- (1) Roboter handeln in der physikalischen Welt, so dass HRI gemeinsames Handeln von Mensch und Roboter sowie die physische Interaktion von Menschen und Robotern umfasst,
- (2) Roboter nehmen die physikalische Welt wahr, so dass HRI die Kommunikation über gemeinsam wahrnehmbare Objekte einschliesst und
- (3) Roboter können durch physische Präsenz, ihre Handlungs- und Kommunikationsweisen sowie gegebenenfalls durch ihre „Körperliche Erscheinung“ ihre Interaktionspartner zu einem anthropomorphen Verhalten beeinflussen.

Mensch-Roboter Interaktion Situiertheit und Grounding

- (1) Erfolgreiche Interaktion zwischen Mensch und Roboter setzt voraus, dass Mensch und Roboter die Entitäten in der Welt in der Kommunikation so spezifizieren, dass der Kommunikationspartner die „richtigen Entitäten identifiziert“.
→ co-reference, grounding, anchoring
- (2) Mensch-Roboter Interaktion ist situiertheit:
 - a) Mensch und Roboter kommunizieren über „räumliche Ausschnitte der Welt“ im Hinblick auf eine Zeitphase
 - b) Kommunikationssituation (K-Sit) und Handlungssituation (H-Sit) müssen nicht räumlich und zeitlich übereinstimmen
 - c) Mensch und Roboter müssen während K-Sit und H-Sit nicht ko-lokalisiert sein
→ displaced reference

Mensch-Roboter Interaktion



Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 – 358

The bottle game:
I want the red bottle!



The bottle game:
I want the small bottle!



The bottle game:
I want the Swiss bottle!



The bottle game:
I want the black bottle!



The bottle game:
I want the black bottle!



The bottle shop
Some other variants



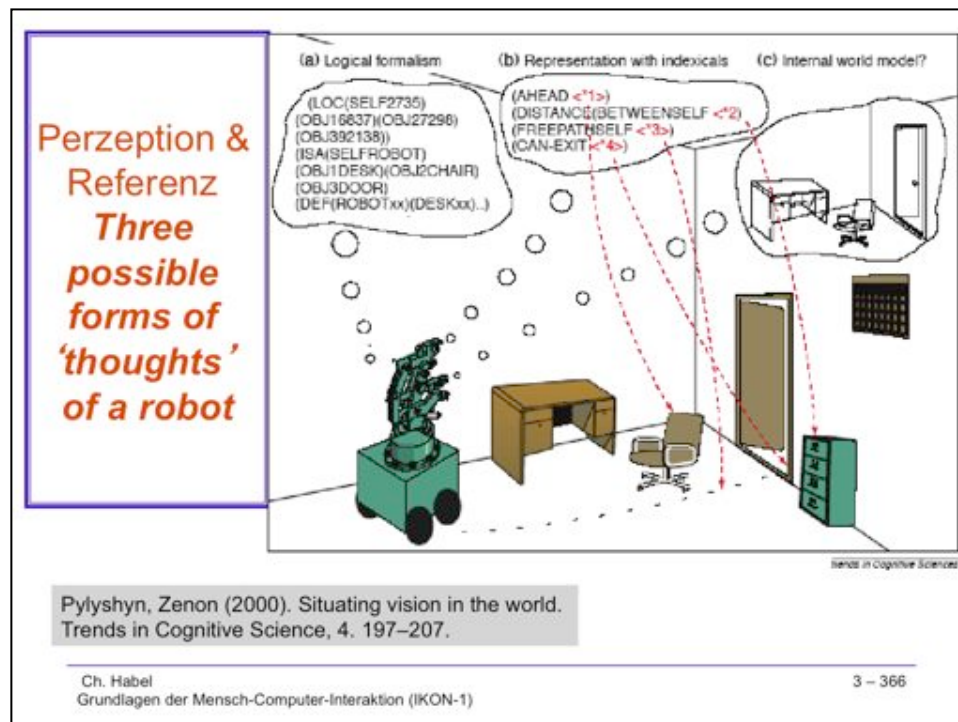
Perzeption & Referenz

Ein Objekt visuell zu erkennen, bedeutet

- verschiedene visuelle Eindrücke und die temporären internen Modelle dieser Eindrücke mit einer Entität in der externen Welt in Verbindung zu setzen
- diese wahrgenommene Entität der externen Welt mit eine interne Repräsentation, d.h. einen internen Stellvertreter, in Beziehung zu setzen
- Etablierung von **Co-Referenz**

Zwei wesentlich verschiedene Fälle

- Erkennung einer individuellen Entität als diese spezifische Entität
- Erkennung einer individuellen Entität als Mitglied einer Kategorie



- Pylyshyn, Zenon (2000). Situating vision in the world. *Trends in Cognitive Science*, 4. 197–207.

Interne Struktur von Symbolen Komplexe Symbole

- Syntax
von Symbolsystemen / Zeichensystemen / Sprachen
 - basiert auf einem Inventar / Vokabular / Alphabet
 - Bestimmt die Menge der wohlgeformten Ausdrücke
 - korrespondiert zur Semantik

→ Eine produktive Syntax ist die Voraussetzung für ein flexibles und leistungsfähiges Kommunikationssystem:

- Grosse, eventuell unendliche, Mengen von Sachverhalten müssen kommuniziert werden können, und zwar mit einem endlichen Inventar von externen Repräsentationen.

- Die informatische Sichtweise komplexer Symbolsysteme wird ausführlich in der Veranstaltung FGI-1 (Formale Grundlagen der Informatik - 1) im 2. Semester der B.Sc. Studiums vermittelt. Dabei wird einerseits eine spezielle Sprache, die der mathematischen Logik behandelt, an der exemplarisch die Prinzipien des Zusammenspiels von Syntax und Semantik erläutert wird. Andererseits wird mit der Theorie formaler Sprachen ein genereller Ansatz für die komplexe Symbolsysteme erläutert.

Komponenten des Sprachsystems

- Grammatik = Wissen über die Sprache
 - Phonologie: Lautsystem
 - Morphologie: Regelsystem der Wortbildung
 - Syntax: Regelsystem der Satzbildung
 - Semantik: System der Bedeutung
- Lexikon
 - System der Basisentitäten der Sprache: *Lexeme*
- Pragmatik
 - System der systematischen Einbeziehung – einschliesslich Vorwissen – in die Bedeutung

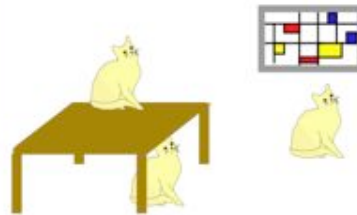
„Grammatik = Wissen über die Sprache“

- ist die kognitionswissenschaftliche Perspektive auf Sprache
- ist insbesondere durch Noam Chomsky in die moderne Linguistik eingeführt worden.

Interne Struktur von Zeichen – Komplexe Zeichen

■ Sätze der natürlichen Sprache:

- *Die Katze liegt auf dem Tisch.*
- *Die Katze sitzt auf dem Tisch.*
- *Die Katze sitzt unter dem Tisch.*
- *Die Katze sitzt unter dem Bild.*
- *Das Kind sitzt unter dem Bild.*
- *Zwei Kinder sitzen unter dem Bild.*



Referenz findet auf alle Ebenen der komplexer Zeichen statt.

- Die Phrase “die Katze” verweist in der gegebenen Sätzen jeweils auf ein Tier einer gewissen Gattung. Die Situation bzw. der Satz spezifizieren, um welches Tier es sich handelt.
- Im Fall des 3. Satzes wird durch “unter dem Tisch” auf eine spezifische Region eines Raumes verwiesen. Hierdurch wird auf die Referenzbeziehung für “die Katze” eindeutig.
- Die Sätze 2 - 4 unterscheiden sich dadurch, dass die räumlichen Präpositionen *auf* und *unter*, bzw. die Nomen innerhalb der Präpositionalphrase, also *Tisch* vs. *Bild*, unterscheiden. Somit liefern die drei Sätze Referenz auf drei verschiedene Regionen.
- Der Satz “Die Katze sitzt auf dem Tisch” verweist auf eine Situation. Dies ist ein Fall, in dem ein komplexe Symbol, nämlich ein Satz, auf eine komplexe Entität, nämlich eine Situation verweist.

Sprachliche Repräsentation & Referenz

▪ Sprachliche Beschreibung der Welt

Die Katze sitzt auf dem Tisch.

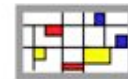
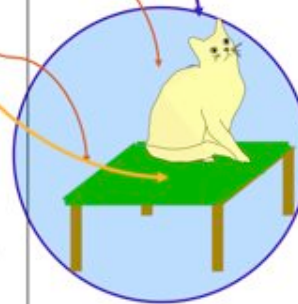
An der Wand hängt ein Mondrian.

▪ Typen von Entitäten

- Physikalische Objekte
- Situationen, Ereignisse
- Regionen
- Zeiten

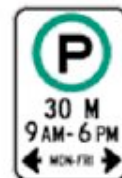
Bis vor zwei Minuten sass die Katze auf dem Tisch.

▪ Die Welt



Interne Struktur von Zeichen – Komplexe Zeichen: Syntax & Semantik von Verkehrszeichen

- bedeutungstragend sind, u.a.,
 - Farben
 - Form
 - Symbole, Bilder
- Bedeutungsinventar
 - *Permission*
blau (DE), grün (US)
 - *Regulation*
kreisförmig (DE)
 - *Prohibition*
diagonaler Balken (DE),
rot (US)



Permission



Prohibition



- Dies ist eine beispielhafte Sammlung von deutschen und US-amerikanischen Verkehrsschildern zum Thema "Parken und Nicht-Parken"

Ein Zwischenfazit: Zeichensysteme – Sprachsysteme

- Zeichensysteme – Sprachsysteme sind veränderbar
 - mit der Zeit
 - durch die Kommunikationsteilnehmer
 - um flexibel auf neue Situationen anwendbar zu sein
- Interne Struktur von Zeichen- / Sprachsystemen
 - Effizienzanforderungen: Verarbeitung & Speicherung
 - Flexibilität: Anpassbarkeit an die Kommunikationssituation
- ➔ Optimierungsdilemma
 - Robustheit – Flexibilität – Exaktheit –
- Natürliche Sprache ist ein erfolgreiches Exemplar eines *optimierten Kommunikationssystems*

Wie einfach ist Sprache?

- Sprache zu produzieren und zu verstehen,
 - erfordert keinen besonderen kognitiven Aufwand,
 - kann – fast – jeder,
- Aber
- der Spracherwerb
 - erfordert einen Zeitraum von mehreren Jahren bei Kindern,
 - ist noch aufwendiger und schwieriger Erwachsenen,
- Sprachverstehen und Sprachproduktion
 - ist maschinellen Systemen bisher nur sehr eingeschränkt möglich.

Wider einige Missverständnisse:

- Sprache kann (fast) jeder, also ist Sprache einfach.
- Sprache lernen schon kleine Kinder, also ist es ganz einfach.
- Wir sprechen ohne es gross planen zu müssen (beinahe automatisch vom Denken zum Sprechen)
- Menschen sind Experten im Sprechen und Sprach verstehen, aber sie wissen meist wenig darüber, wie dies - im Detail - abläuft.
- Die wissenschaftliche Untersuchung sprachlicher Prozesse, ist ein äusserst komplexer Bereich interdisziplinärer Wissenschaft:
 - Linguistik, Informatik (Computerlinguistik), Psychologie, Neurowissenschaft

„Why can ‘t computers use English?“ [Ray Jackendoff]

- Szenario: Die Vorlesemaschine
als Beispiel für Probleme bei der maschinellen
Sprachbeherrschung.
Lesen von – englischem – Text.
- *read* mögliche Aussprache: [ri:d], [red]
(1) The girls will read the paper. [ri:d]
(2) The girls have read the paper. [red]
→ ein erster Vorschlag:
 - direkt nach *will* Aussprache [ri:d]
 - direkt nach *have* Aussprache [red]

- Ray Jackendoff, “Why can’t computers use English”. published by
The Linguistic Society of America.
http://www.lsadc.org/info/pdf_files/Why_can't_computers.pdf
(URL vom 12.12.2008)
- Vorlesemaschine: Text-to-speech
(etwa für Blinde) ist ein wünschenswertes Softwareprodukt, das aber
bisher nicht in wirklich befriedigender Qualität erstellt werden konnte.
- Text-to-speech Produktion ist hochgradig sprachspezifisch, nicht nur
im Bereich der Intonation
 - Einfache Lösung: Statistische Verfahren

„Why can ‘t computers use English?“ (2)

(3) Will the girls read the paper? [ri:d]

(4) Have the executors of the will read the paper? [red]

- *will* ist nicht direkt vor *read*, trotzdem wird [ri:d] gelesen.
- *will* ist direkt vor *read*, trotzdem wird [red] gelesen.

→ Syntaktische Information wird benötigt:

- *will* und *have* sind Hilfsverben, die das Verb *read* modifizieren. Unterschiedliche Modifikationen führen zu unterschiedlicher Aussprache.
- *will* kann auch ein Nomen sein, dann kann das Wort *will* auch nicht als Hilfsverb die Aussprache von *read* beeinflussen.

„Why can ‘t computers use English?“ (3)

(5) Have the girls who will be in vacation next week read the paper yet? [red]

- Hauptsatz: *have* modifiziert *read*, eingebetteter Relativsatz *will* modifiziert *be*.

(6) Please have the girls read the paper. [ri:d]

- *have* ist hier kein Hilfsverb, sondern ein Vollverb.

(7) Have the girls read the paper? [red]

- *have* ist hier wieder Hilfsverb.

→ Auch / schon sehr einfache Aufgaben der Sprachverarbeitung benötigen syntaktische Analysen.

Gesprochene Sprache – geschriebene Sprache Speech – Language

	<i>speech</i>	<i>language</i>
zeitliche Charakteristik	Produktionszeit = Rezeptionszeit transient, non-persistent	Produktion und Rezeption entkoppelbar persistent / dauerhaft
Realisierung	akustisch, in kontinuierlichem Medium	visuell, basierend auf diskreten Symbolen
Universalität & Evolution	„ursprüngliche“ Form der Sprache“, Grundlage (fast) aller Sprachen	historisch junge Entwicklung; existiert nicht für alle Sprachen
Erlernbarkeit	wird – weitgehend automatisch während der frühen Kindheit erworben	wird durch Training auf der Basis gesprochener Sprache erlernt.
Braille: Alphabet für die haptische Wahrnehmung von geschriebener Sprache „Gestensprachen“ sind nicht gestische Varianten gesprochener Sprache		

Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 – 377

- Sprachliche Äusserungen sind externe Repräsentationen

Multimodalität		
Gesprochene Sprache – geschriebene Sprache		
	<i>speech</i>	<i>language</i>
	Verarbeitung von visueller Perzeption (weitgehend) entkoppelt (Multimodale Informationsverarbeitung) McGurk effect	Verarbeitung mit visueller Perzeption verbunden (→Interaktionen im Arbeitsgedächtnis)

Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 – 378

McGurk Effect

- McGurk, Harry & MacDonald, John (1976). Hearing lips and seeing voices. *Nature*, 264. 746–748.
- The McGurk Effect refers to an auditory illusion in which visual cues to the syllable "ga" are combined with auditory cues to syllable "ba" resulting in the perception of "da" or "tha". Try listening to the following video with your eyes closed, then, after several repetitions, open your eyes to see how your perception changes in the presence of the visual stimulus.
- Die Web-Site von Patricia Kuhl, der auch die obige Charakterisierung entnommen ist, gibt Ihnen die Möglichkeit den Effekt zu erleben.
- <http://ilabs.washington.edu/kuhl/research.html>

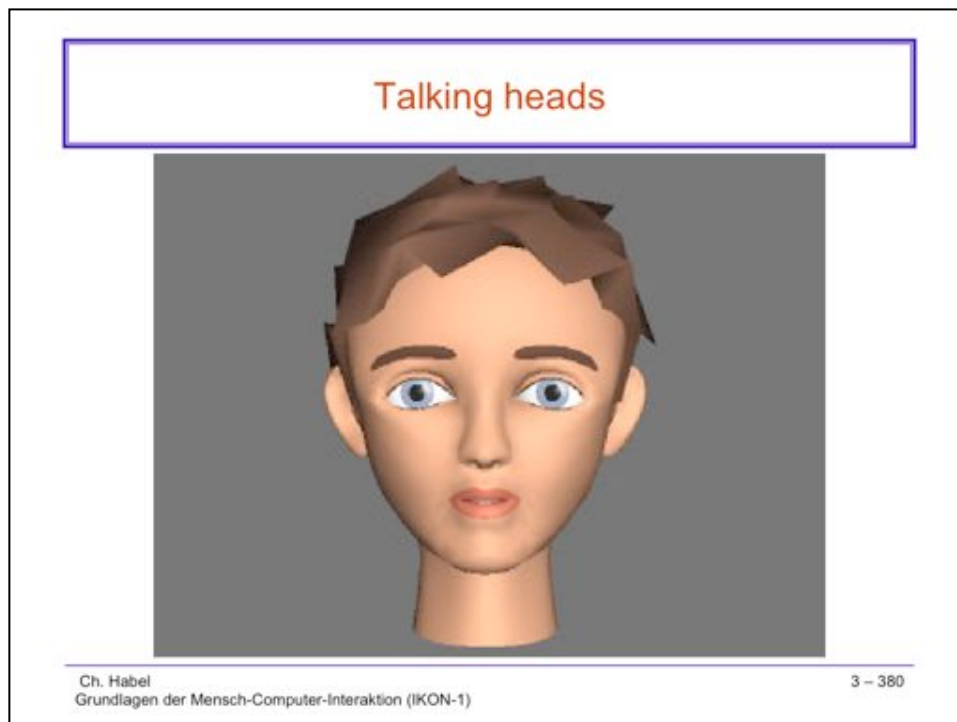
McGurk effect



Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 – 379

- **McGurk Effect:**
- The McGurk Effect refers to an auditory illusion in which visual cues to the syllable "ga" are combined with auditory cues to syllable "ba" resulting in the perception of "da" or "tha". Try listening to the following video with your eyes closed, then, after several repetitions, open your eyes to see how your perception changes in the presence of the visual stimulus.
- <http://ilabs.washington.edu/kuhl/research.html>



- “Talking Heads” ist die – im HCI-Bereich übliche – Bezeichnung für *animierte / künstliche Agenten*,
 - von denen nur der Kopf in der Visualisierung erscheint,
 - die mit dem Benutzer sprachlich interagieren, und zwar durch gesprochene Sprache,
 - und die über *Gesichtsmimik* verfügen.
- Die Entwicklung von *Talking heads* basiert auf einer Integration von Visualisierungstechniken und Verfahren aus dem Bereich *speech synthesis / speech production*.
- Die Verständlichkeit eines *talking heads* – und hieraus folgend die Akzeptanz durch den Benutzer – hängt von der Modellierung der Interaktion von Mimik & Sprache ab.
 - Nicht natürliche Visualisierung der Mundpartie / der Lippenbewegungen kann dazu führen, dass die Erkennung gesprochener Sprache bei synchronen Darbietung eines *talking heads* schwieriger bzw. mit mehr Fehlern verbunden ist.
 - Ebenso erschweren Ungenauigkeiten in der zeitlichen Synchronisation zwischen Sprechmimik und Sprachproduktion das Verständnis der künstlichen Sprache durch den Benutzer.
- Der auf dieser Folie dargestellte *Talking head* RUTH wurde durch Matthew Stone und seine Arbeitsgruppe an der Rutgers University entwickelt. Animierte Versionen (mit sprachlichen Äusserungen können über die URLs <http://www.cs.rutgers.edu/~mdstone/> und <http://www.cs.rutgers.edu/~village/ruth> erreicht werden.

Das Original



Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 - 381

Multimodale Kommunikation durch Sprache und bildhafte Repräsentationen

- Sprache
 - gesprochen vs. geschrieben
- Bildhafte Repräsentationen
 - Informationsgraphiken
 - Karten
- Ziele
 - Kombination der Stärken verschiedener repräsentationeller Modalitäten
 - Ergänzung bei Verarbeitung von externen Repräsentationen im Rahmen sensorischer Substitution

In einem früheren Abschnitt wurde der Unterschied zwischen sensorischen und repräsentationellen Modalitäten angesprochen. Dieser Sichtweise folgend sind *Sprache* und *Bildhafte Repräsentationen* verschiedene repräsentationelle Modalitäten, die in der Kommunikation eingesetzt werden können. [betrifft gleicher Weise die Kommunikation zwischen Menschen und zwischen Menschen und Computern.

- Innerhalb dieser repräsentationellen Modalitäten gibt es Submodalitäten, z.B. gesprochene vs. geschriebene Sprache.
- Im folgenden werden exemplarisch zwei Bereiche von multimodaler Kommunikation in der Kombination von Sprache und bildhaften Repräsentationen vorgestellt.

Referenz- und Koreferenz-Beziehungen in multimodal Dokumenten

Abb. 1 die jährliche Abweichung der August-Temperatur vom langjährigen Durchschnitt (Norm 1961-1990).

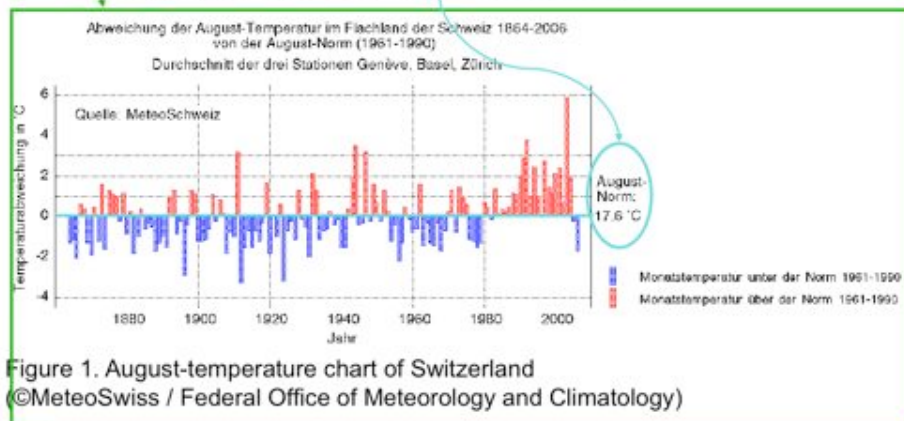


Figure 1. August-temperature chart of Switzerland
(©MeteoSwiss / Federal Office of Meteorology and Climatology)

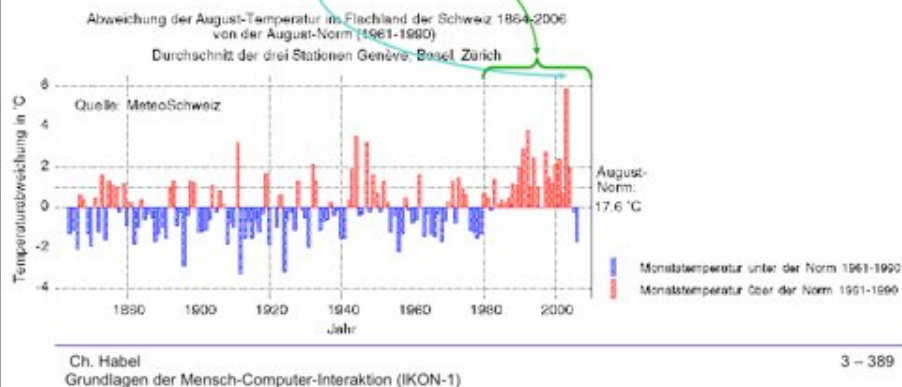
Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 – 388

- Die Kombination von Texten und Informationsgraphiken ist in Fachtexten aber auch in der journalistischen Information (Zeitungen, Internet, Fernsehen) von grosser Bedeutung. Für die Informatik ergeben sich hieraus u.a. die folgenden Anwendungsfelder:
 - Information retrieval: Bei der Recherche von multimodalen Dokumenten sollte auch die Information, die in Graphiken repräsentiert wird, auffindbar sein, und zwar in einer Weise, dass die Beziehung zwischen graphisch repräsentierten Inhalten und textuell repräsentierten Inhalten berücksichtigt wird.
 - Zugang zu graphisch repräsentierten Inhalten für Blinde (und Sehbehinderte): Die sprachliche Präsentation von Inhalten, die graphisch repräsentiert im Internet vorliegen, ist eine Möglichkeit, diese Inhalte Blinden zugänglich zu machen. Dies sollte in integrierter Form geschehen, also derart, dass die aus der Graphik stammenden Inhalte in textuell repräsentierte Inhalte integriert werden (und nicht nach dem Text präsentiert werden).
 - Fortsetzung des Kommentars in den Notizen zur nächsten Folie

Referenz- und Koreferenz-Beziehungen in multimodal Dokumenten

- Die wärmere Periode, die in den 80ern begann, kulminiert in der Spitze von 2003.

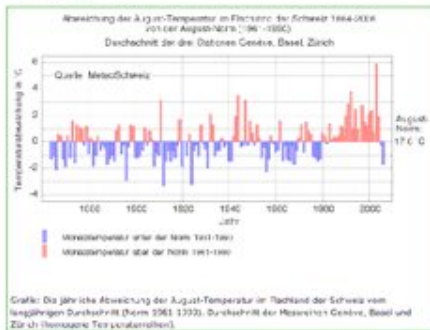


- Die zentrale Voraussetzung für die Berechnung der integrierten Bedeutung (aus den beiden Modalitäten) ist das Auffinden von Koreferenzbeziehungen, d.h. von Konstellation, in mehrere repräsentationelle Entitäten auf das gleiche Objekt in der Welt verweisen:
 - Zu Referenz:
 - Die Graphiken und die meisten der deutschsprachige Texte, die diesen drei Folien zur Klimaveränderung diskutiert werden, stammen aus einer Mitteilung von MeteoSwiss
http://www.meteoschweiz.ch/web/de/wetter/wetterereignisse/kalter_august_06.html
 - Eine ausführliche Diskussion aus der Perspektive multimodaler Dokumente gibt:

Habel, Christopher & Acartürk, Cengiz (2007). On reciprocal improvement in multimodal generation: Co-reference by text and information graphics. In I. van der Sluis, M. Theune, E. Reiter & E. Krahmer (eds.) Proceedings of the Workshop on Multimodal Output Generation (MOG 2007). 25. – 26. January, 2007. Aberdeen, United Kingdom. 69–80.

Interne Struktur eines Text-Graphik Dokuments

01. September 2006 / Stephan Bader
Wie kalt war der August 2006 wirklich?



Verändertes August-Klima
Gegen Ende der 1990er Jahre haben sich die August-Temperaturen auffallend verändert, indem sie sich spürbar um mehr als 1 Grad erhöhten. Gekühlt vor 1990, kühlte August. Temperatur, wie die oben veränderte, ist ein Hinweis auf den Klimawandel. In diesem Sinne ist der kalte August 2006 nichts anderes als die Wiederaufnahme der kalten Tradition, und somit eigentlich nichts ungewöhnliches.

Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

Überschrift

Abbildung

Titel

- Legende
- Framework
 - outer vs. inner
 - Markierungen und Beschriftungen

➤ Kombination von bildhaften und textuellen Elementen

Text

3 – 390

Karten für Blinde

(eine Vorschau auf die zu lösenden Aufgaben)

- Sensorische Substitution
Tactile / haptische Wahrnehmung an Stelle von visueller Wahrnehmung
 - taktile Karten
 - taktilen Kartenlesen – Design taktiler Karten
- Gebrauch von taktilen Karten in der Navigation
 - Selbstlokalisierung
 - Aufbau mentaler Karten
- Ergänzung von taktilen Karten durch verbal präsentierte Information

- Als zweiter Anwendungsbereich von multimodaler Kommunikation, in der Sprache und graphische Repräsentation interagieren, wird die Nutzung von Karten durch Blinde oder stark Sehbehinderte vorgestellt..
- Diese Thematik ist Gegenstand aktueller Forschung (seit 2006/7) am Arbeitsbereich WSV des Departments Informatik
- In der Vorlesung folgt eine Demonstration der visuell-sequentiellen Exploration einer Karte (analog zu einer entsprechenden Demonstration in Abschnitt zur Haptischen Wahrnehmung).

Sequentielle Wahrnehmung einer Karte

- Gleich sehen Sie – wieder durch eine Lochblende – eine Karte.

➤ Versuchen Sie **ANSCHLIESSEND** eine Skizze zu zeichnen.

- Kindred to experiments of
- Loomis, J. M., Klatzky, R. L., & Lederman, S. J. (1991) Similarity of tactual and visual picture perception with limited field of view. *Perception*, 20, 167-177. [loomis_91.pdf]

Pfad Skelett für die taktile Wahrnehmung generiert aus einer 1:5000 topographischen Karte



Die durch die Lochblende wahrgenommene Route

Ch. Habel
Grundlagen der Mensch-Computer-Interaktion (IKON-1)

3 – 423

- Die rechte Abbildung zeigt einen Ausschnitt des 'Katasterblattes' (DK5 / Massstab 1:5000) der Landesbetriebs für Geoinformation und Vermessung.

<http://www.hamburg.de/geobasisdaten/367904/dk5.html>

- Die linke Abbildung hebt das Strassennetz hervor. Die in der Vorlesung – durch die Lochblende – präsentierte Route ist gelb unterlegt.

Ergänzung taktiler Karten durch Sprache

- Taktile Karte + Sprache
 - Gedruckte Karten: Dialoge zwischen blinden und sehenden Personen
 - visuo-tactile maps
 1. Trainingskontext: Erwerb der Fähigkeit Karten – taktil – zu lesen
 2. Navigationskontext: verbal unterstützter Erwerb von räumlichen Wissen aus Karten
 - force-feedback maps: automatisch Erzeugung gesprochener Sprache im Trainings oder Navigationskontext
 - audio-tactile maps

- Da die ausschliesslich taktile Wahrnehmung von Karten eine hohe kognitive Belastung darstellt (vgl. Haptische Wahrnehmung zu Problemen sequentieller visueller und taktiler Wahrnehmung und Integration der -zeitlich-sequentiellen räumlichen Eindrücke zu einem räumlichen Gesamteindruck), bietet sich die Unterstützung der taktilen Wahrnehmung durch sprachlich kodierte Information an.
- Dies wird in den weiteren Folien durch auditiv wahrnehmbare Kommentare zu verschiedenen Stadien der taktilen Exploration erläutert.

Anwendungen der Computerlinguistik

- Mensch–Computer Interaktion
 - gesprochene / geschriebene Eingabe & Ausgabe
 - allgemeine Prinzipien der Kommunikation
- Assistenz für die menschliche Kommunikation
 - maschinelle Übersetzung
 - Text-to-speech / speech-to-text
 - Sprachlehr- / lernsysteme
 - Unterstützung von Personen mit Behinderungen
- Maschinelle Verarbeitung von Texten
 - Informationsextraktion aus Texten (WWW)
 - Information Retrieval
 - Filterung und Verteilung von Texten

Das IAF Computerlinguistik im B.Sc. Informatik

- Linguistisches Modul
 - Einführung in die Linguistik
 - besteht aus Vorlesung und Seminaren
 - wird im Wintersemester durchgeführt
 - wird in verschiedenen einzelsprachlichen Disziplinen angeboten:
für CL ist „Linguistik des Deutschen“ präferiert.
- Informatisches Modul
 - Grundlagen der Computerlinguistik
 - betrifft: Datenstrukturen und algorithmische Verfahren zur
Bearbeitung natürlicher Sprache, z.B. Syntaxanalyse,
Bedeutungszuweisung
 - basiert auf:
 - Grundlagen der theoretischen & praktischen Informatik
 - Grundkonzepten der Linguistik

Warum sollten Studierende der Informatik sich in Computerlinguistik vertiefen ?

- Einblick in eine / bzw. mehrere andere Disziplin(en)
 - Erweiterung des wissenschaftlichen Horizonts
 - Vorbereitung auf die Praxis: InformatikerInnen haben es im Beruf (fast) immer mit Nicht-InformatikerInnen zu tun.
- Aufbau von Kenntnissen in einem wichtigen Bereich der Informationstechnologie
 - Sprachtechnologie ist eine der Basistechnologien des Web-Zeitalters
 - Sprachtechnologie hat starke einzelsprachliche Anteile:
„Die Sprachsysteme des Deutschen werden (fast ausschliesslich) in Deutschland entwickelt werden.“
- Allgemeinbildung: Verbesserung der sprachlichen Kompetenz.

Integrierte Anwendungsfächer

- Die Idee:
 - Kombination von Informatik und einem anderen Fach
 - abgestimmte Inhalte mit – im wesentlichen – gleichem Umfang (Summe 12 SWS)
 - gemeinsame Zielsetzung
 - Anwendung auf der Basis der Integration von Informatik und dem anderen Fach
- Die Durchführung:
 - Gemeinsame oder abgestimmte Veranstaltungen
 - Zwei Moduln: In der Regel im 5. und 6. Semester