

Grundlagen der Kognitiven Informatik

Einführung

Ute Schmid unterstützt von Michael Siebers

Kognitive Systeme, Angewandte Informatik, Universität Bamberg

letzte Änderung: 27. Oktober 2010

Organisatorisches

- Modul mit 3 ECTS für BA-Psychologie (ab 5. Semester) als einer von zwei Kursen im Wahlpflichtbereich Kognitionswissenschaften
- Als gute Vorbereitung für eine Vertiefung im Bereich Informatik im Master
- Modul ist als V + Ü konzipiert (etwa 50% Vorlesung, etwa 50% Übungen, zum Teil Übungsaufgaben)
- Prüfung: Hausarbeit
- Der Kurs ist ebenfalls offen für Nebenfachstudierende (30 oder 45 ECTS Nebenfach)

Literatur/Quellen

- Schmid & Kindsmüller (1996), Kognitive Modellierung, Spektrum
- Sun (Ed., 2008), The Cambridge Handbook of Computational Psychology, CUP.
- Dörner & Schmid (in Druck), Computermodellierung psychischer Prozesse, in Schütz et al., Psychologie (4. Auflage), Kohlhammer

Organisationen

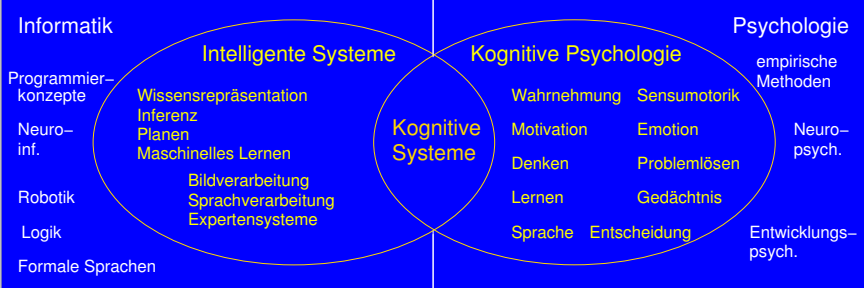
- Studentische Mitgliedschaft in der Gesellschaft für Kognitionswissenschaft (www.gk-ev.de)
- Cognitive Science Society
wichtige Zeitschriften: *Cognitive Science*, *Trends in Cognitive Science* (TiCS)
- Wichtige Tagungen: Fachtagung der Gesellschaft für Kognitionswissenschaft (KogWis 2012 in Bamberg!)
- Jährliche Konferenz der Cognitive Science Society

Was ist Kognitive Informatik

- Informatik, die Erkenntnisse der (kognitiven) Psychologie bei der Entwicklung und Gestaltung von Algorithmen und interaktiven Systemen berücksichtigt („Psychonik“)
↔ ingenieurwissenschaftliche Perspektive: performanz-/erfolgsorientiert)
- Formalisierung, Präzisierung, Implementierung von Theorien der (kognitiven) Psychologie sowie Analyse genereller Beschränkungen (**kognitive Modellierung**, generative Theorien, Baupläne)
↔ erkenntniswissenschaftliche Perspektive: Werkzeug der theoretischen Psychologie)

siehe Schmid, 2008, KI Themenheft Kognition

Einordnung: Kognitive Modellierung



Was ist ein Modell?

- Modelle werden gebaut, um komplexe Objekte oder Systeme besser zu verstehen
- Modelle erfassen relevante Aspekte eines Gegenstandsbereichs und abstrahieren von anderen Aspekten
- Arten von Modellen:
 - ▶ Physikalisches Modell (Architektur, Maschinenbau, Medizin)
 - ▶ Mathematisches Modell (Meteorologie, Psychologie, Physik)
 - ▶ Objektorientiertes Modell (Software-Entwicklung)
 - ▶ Computationales Modell (Astronomie, Volkswirtschaft, Psychologie)

Ein kognitives Modell

ist ein Computerprogramm

- dessen Verhalten in einer bestimmten Beziehung ähnlich zu menschlichem Verhalten ist
- durch dessen Entwicklung und Nutzung wir erhoffen, Einsicht in Strukturen und Prozesse menschlichen Denkens und Handelns zu erhalten

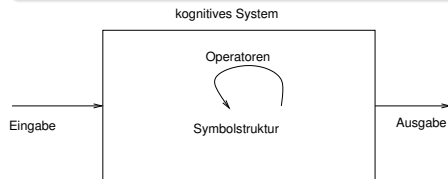
zu klären:

- *Ähnlichkeit zu menschlichem Verhalten*

Informationsverarbeitungsansatz

Grundannahme bei kognitiven Modellen:

- Eingaben werden intern mittels mentaler Operatoren zu Ausgaben transformiert
- Die Black Box des Behaviorismus wird durch Annahmen über Informationsverarbeitungsprozesse ersetzt
- Von Interesse ist die „geistige“ Ebene, nicht die neurophysiologische Ebene
- Metapher: Geist zu Gehirn ist wie Software zu Hardware (Marvin Minsky)
- \leftrightarrow Modellierung als Computerprogramm möglich



Beispiel: Schlussfolgerndes Denken

- Aufgabe: gegeben eine Menge von Fakten der Form $relation(x, y)$ soll beurteilt werden ob Aussagen der Form $relation(a, b)$ korrekt sind
- Abhängige Variable: Reaktionszeit (bei korrekter Antwort)
- Zwei kognitive Modelle:
 - ▶ Logische Schlussfolgerung
 - ▶ Konstruktion einer integrierten, relationalen Repräsentation aller Prämissen, „Ablese“ der Aussage (Symbol-Distanz-Effekt, Potts, 1975)

Beispiel

```
taller(Anton, Bernd).  
taller(Bernd, Claus).  
taller(Claus, Dieter).  
taller(Dieter, Emil).
```

Modell 1: Transitivitätsregel

```
isTaller(X,Y) :- taller(X,Y).  
isTaller(X,Y) :- taller(X,Z), isTaller(Z,Y).
```

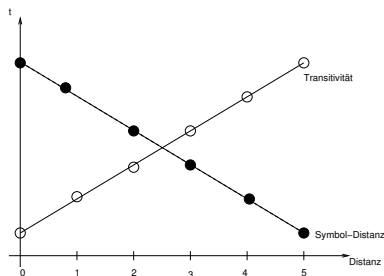
Vorhersage: Antwortzeiten für gegebene Fakten am kürzesten, je öfter Transitivitätsregel angewendet werden muss, desto länger die Antwortzeit

Modell 2: Integrierte Repräsentation

Anton	Bernd	Claus	Dieter	Emil
<hr/>				
<i>tallest</i>				<i>smallest</i>

Vorhersage: (Symbol-Distanz-Effekt, ...) Je weiter entfernt zwei Objekte in der mentalen Repräsentation sind, desto kürzer ist die Antwortzeit

Symbol-Distanz-Effekt



- Vergleich zweier unterschiedlicher Annahmen über kognitive Prozesse via Vorhersage von Reaktionszeitunterschieden
- Beispiel für eine sinnvolle Möglichkeit, um das Verhalten eines kognitiven Modells auf Ähnlichkeit zum menschlichen Verhalten zu überprüfen

Ergänzende Bemerkungen

- vgl. Theorie der Mentalen Modelle von Johnson-Laird, siehe z.B. chap. 12 in Handbook of Computational Psychology
- Reihenfolge der Prämissen kann Aufbau einer integrierten Repräsentation beeinflussen
- weiteres Beispiel: Vorhersagen des Teachable Language Comprehender (TLC, Quillan, 1969)

Computermodelle menschlicher Informationsverarbeitung

- Kognition betrifft *mentale* Aktivitäten
- Keine *direkte* Beobachtbarkeit via Eingabe-/Ausgabe-Verhalten
- Vorannahme: Informationsverarbeitungsansatz, Symbolsysteme
- Empirischer Zugang: Hypothesen über kognitive Prozesse, die zu spezifischen Annahmen führen, wie verschiedene Aufgaben sich auf Reaktions-/Lösungszeiten oder andere abhängige Maße (wie Fehler) auswirken
- Modellierung: Werkzeug um vollständige und konsistente ablauffähige Modelle über kognitive Prozesse zu erzielen
- Starke vs. schwache KI

Charakteristiken guter Theorien

- konsistent (intern und extern)
- sparsam (möglichst wenig theoretische Konzepte und Erklärungsmuster)
- nützlich (beschreibt und erklärt beobachtbare Phänomene)
- empirisch testbar und falsifizierbar
- gestützt durch kontrollierte, wiederholbare Experimente
- korrigierbar und dynamisch (Änderungen aufgrund neuer Befunde)
- progressiv (Erweitert Geltungsbereich vorhandener Theorien)
- tentativ (räumt ein, dass Annahmen nicht korrekt sein könnten)

↔ kognitive Modellierung hilft, Theorien auf Widerspruchsfreiheit und Vollständigkeit zu prüfen

Computational Mind?

Was ist der Status eines kognitiven Modells?

Is the mind a computational phenomenon? No one knows. It may be; or it may depend on operations that cannot be captured by any sort of computer. ...

Theories of the mind, however, should not be confused with the mind itself, any more than theories about the weather should be confused with rain or sunshine. And what is clear is that computability provides an appropriate conceptual apparatus for theories of the mind.

This apparatus takes nothing for granted that is not obvious. ... any clear and explicit account of, say, how people recognize faces, reason deductively, create new ideas or control skilled actions can always be modelled by a computer program. (Johnson-Laird, The Computer and the Mind, 1988)

Grundannahmen des Informationsverarbeitungsansatzes

(nach Strube, 1993)

- Unterschiedliche Prozesse basieren auf gemeinsamen, allgemeinen, grundlegenden Prinzipien
- Kognition lässt sich unabhängig von der Ebene ihrer materiellen Basis betrachten
- Informationsverarbeitungsprozesse lassen sich als formale Prozesse beschreiben
- Informationsverarbeitende Prozesse operieren auf Repräsentationen

Anmerkungen zum Informationsverarbeitungsansatz

- Church-Turing These: Jede berechenbare Funktion ist als Turing-Maschine (in einer allgemeinen Programmiersprache) berechenbar.
- Problem: Ein und dasselbe Ein-/Ausgabeverhalten kann durch beliebig viele Computerprogramme erzeugt werden.
- Umsetzung in ein Programm hat zunächst nur den Status einer Machbarkeitsdemonstration. Weitere Restriktionen für kognitive Modelle sind notwendig (z.B. gemeinsame, grundlegende Vorannahmen, siehe kognitive Architekturen)
- Erste Formulierung von Vorannahmen im Rahmen der *physical symbol systems hypothesis* (Newell & Simon, 1963)

PSS Hypothesis

- A physical symbol system (also called a formal system) takes physical patterns (symbols), combining them into structures (expressions) and manipulating them (using processes) to produce new expressions.

PSS Hypothesis: *A physical symbol system has the necessary and sufficient means of general intelligent action.*

- Sehr starke Behauptung/starke KI (bauen eines intelligenten Systems vs. Modellieren eines intelligenten Systems)
- Menschliches Denken ist Symbolmanipulation (PSS ist notwendig)
- Maschinen können intelligent sein (PSS ist hinreichend)

Beispiel: Symbolstruktur

Elementare Symbole $S = \{0, 1, \epsilon\}$

Aufbau von Ausdrücken (Symbolstrukturen) durch Regeln:

$$R_1 : E \rightarrow \epsilon$$

$$R_2 : E \rightarrow 0E$$

$$R_3 : E \rightarrow 1E$$

(Definition einer formalen Sprache)

Beispiel: Ableitung von 100ϵ mit R_3, R_2, R_2, R_1

Beispiel: Transformation von Symbolstrukturen

(Kognition als Berechnung)

Kaffee-Bohnen-Problem:

Gegeben ist eine Kaffeedose, in der schwarze (S) und weiße (W) Bohnen in einer festen Reihenfolge angeordnet sind,

beispielsweise: *W W S S W W S S*.

Gegeben sind folgende Regeln:

$S W \rightarrow S$

$W S \rightarrow S$

$S S \rightarrow W$

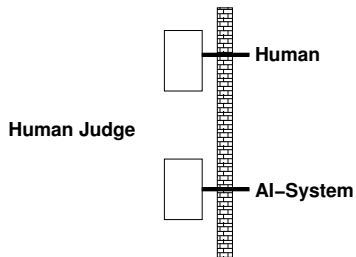
Das Ziel ist, am Ende möglichst wenige Bohnen zu haben. Die Konfliktlösungs-Strategie sei, immer die oberste anwendbare Regel auszuwählen.

W W S S W W S S
W W S S W S S
W W S S S S
W S S S S
S S S S
W S S
S S
W

Kriterien für Kognitive Modelle

- Spezifikation der Anforderung „Ähnlichkeit zu menschlichem Verhalten “
- Verhaltensähnlichkeit:
 - ▶ nur beobachtbares Verhalten
 - ▶ Vergleiche: Turing-Test
 - ▶ Ein und dasselbe Eingabe-/Ausgabe-Verhalten kann durch unendliche viele verschiedene Programme realisiert werden!
- Berücksichtigung interner Zustände und Prozesse als Constraints!
 - ▶ Zeitverlauf (relativ, absolut)
 - ▶ Erfolg/Fehler
 - ▶ Relative Aufgabenschwierigkeit

Einschub: Turing Test



- Kasparov vs. Deep Blue, Match 1996: Kasparov vermutete menschliche Intervantion
- Searle Kritik: Gedankenexperiment "Chinese Room"
- Vergleich Menschliche Kognition und Modell ist immer nur *functional* nicht *strukturell* möglich

Ansätze der Kognitiven Modellierung

- Kognitive Architekturen (am bekanntesten)
- Spezielle Modelle für spezielle Aspekte
- Konnektionistische Modelle
- Probabilistische (Bayes'sche) Modelle

Kognitive Architekturen

- Fokus auf Mechanismen der kognitiven Verarbeitung
- Vorgabe von allgemeinen Prinzipien, “unified theory of cognition”
 - ▶ Explizite Definition von Grundmechanismen der Informationsverarbeitung: Kontrolle der Interaktion mit der Umgebung, Repräsentation von Information im Gedächtnis, Strategien zur Regelauswahl
 - ▶ Annahme, dass diese Mechanismen über alle Bereiche konstant sind (Problemlösen, Schlussfolgern, Sprachverstehen, Mustererkennung, etc.)
 - ▶ Vorteil: Verschiedene Modelle, die in der gleichen Architektur realisiert sind, sind vergleichbar
 - ▶ Vision: Zusammenschalten von Modellen über verschiedene Bereiche führt zur Annäherung an komplexe Kognition
 - ▶ Nachteile: Modellierungssprache beschränkt die Art, in der Modelle formulierbar sind, Modelle werden unnatürlich und unelegant
- Beispiele: ACT-R, Soar, Cogent, Clarion, Ikarus, Psi

Spezielle Modelle

- Modellannahmen für einen spezifischen Bereich werden in einer Programmiersprache umgesetzt (Prolog, ML, Matlab, ...)
- Beispiele: HAM (Human associative memory, Anderson), TLC (hierarchische Wissensorganisation, Quinlan), frühe Entscheidungsbaum-Algorithmen, Analogie-Löser von Evans (1968)

Konnektionistische Modelle

- Fokus auf basalen Prozessen (Lernen, Aufmerksamkeit, Wahrnehmung)
- Implementiert als künstliches neuronales Netz, Backpropagation
- PDP von McClelland und Rumelhart führte zur Wiederbelebung von neuronalen Netzwerkmodellen

Probabilistische Modelle

- Aktuelle Arbeiten zum Erwerb von Konzepten und Regeln
- häufig Bayes-Ansatz
- Beispiel Tenenbaum

Grundlagen Kognitiver Modellierung

Bestandteile eines kognitiven Modelles

- Wissensrepräsentation \leftrightarrow Repräsentationssprache (Logik, Semantische Netze, Schemata, ...)
- davon abhängig Wissensverarbeitung \leftrightarrow syntaktische Berechnungsvorschriften (Suchalgorithmen, Produktionsregeln, Deduktionsalgorithmen, ...).

In der Veranstaltung werden wir uns zunächst mit Berechnungsvorschriften/Algorithmen beschäftigen.

- Problemlösen und Suche
- Heuristiken
- Mittel-Ziel-Analyse
- Produktionssysteme