Cognitive Modeling

Memory and Knowledge Representation

Ute Schmid

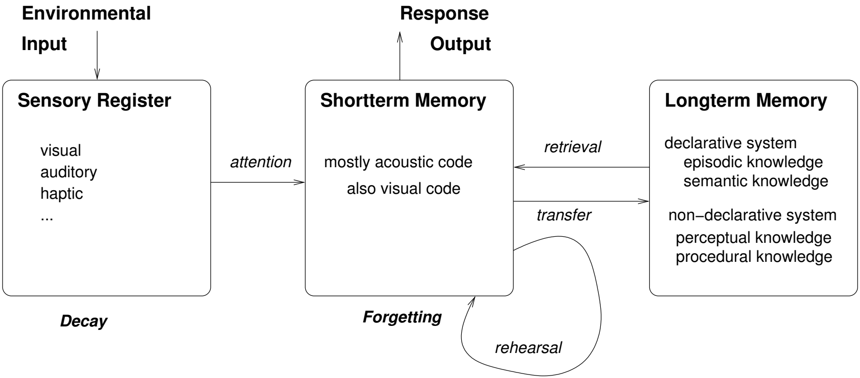
Kognitive Systeme, Angewandte Informatik, Universität Bamberg

last change: 17. Dezember 2014

Human Memory

- Ability to store, retain, and recall information and experiences
- Processes
  - **Sensory memory:** initial 200–500 milliseconds after an item is perceived; no organisation – iconographic; fast decay; Sperling (1960) – partial report paradigm
  - **Short-term memory:** recall for a period of several seconds to a minute without rehearsal; rely mostly on an acoustic code for storing information; limited capacity; can be extended by chunking; George A. Miller (1956) – the magical number 7 plus minus 2
  - **Long-term memory:** very large capacity, potentially unlimited duration; semantic encoding

Modal Memory Model



(Atkinson & Shiffrin, 1968)

Sperling's Span of Apprehension Experiments

First Experiment

- Tachistoscopical presentation of arrays of 12 letters (3 rows of 4 letters each)
- Subjects typically stated that they saw all 12 letters, but could only report 3 or 4 of them before the memory trace faded.

Second Experiment

- Subjects were told that, after seeing the array, they would hear a musical tone (pitched high, medium, or low) telling them which row of the array to report, and the time lapse between the presentation of the array and the onset of the tone was varied as the independent variable.
- With immediate onset (0 sec. delay), subjects can typically report all 4 letters of the indicated row, but with a delayed onset of 1 sec., recall worsens to about 1-2 letters

Sperling's Span of Apprehension Experiments

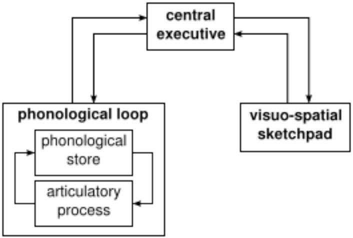
- Limit in the amount of information that can be transferred from sensory store to STM in a given amount of time; due to decay, information not transferred in time is lost!
- Attention processes can govern what information is transferred (and noticed) – Broadbent's filter theory (1958)

Miller's Magical Number 7 plus minus 2

- Modern estimate: 4-5 items
- Chunking can increase memory capacity
- e.g., chunking a 10 digit telefon number into groups
- Chase and Simon (1973): chess experts can remember board constellations not because of a better memory but because they chunk constellations into meaningful units

Working Memory

- Baddeley & Hitch (1974)
- replaced the concept of general short term memory with specific, active components
- Three basic stores:
  - central executive (cp. production system model)
  - phonological loop
  - visuo-spatial sketchpad
- expanded with the multimodal episodic buffer (2000)



Mental Models

- Representations in the mind of real or imaginary situations
- Kenneth Craik (1943): mind constructs "small-scale models" of reality to anticipate events
- constructed from perception, imagination, or the comprehension of discourse
- can be underlying an visual image or be abstract
- are akin to architects' models or to physicists' diagrams in that their structure is analogous to the structure of the situation that they represent
- unlike the structure of logical forms used in formal rule theories

Three Turtles

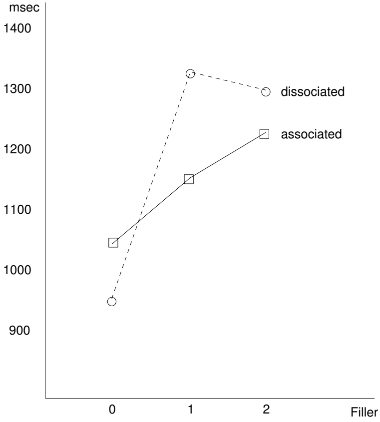
- Bransford, Barclay and Franks, 1972
- Presentation "Three turtles rested on / beside a floating log and a fish swam beneath them."
- Recognition task: "Three turtles rested on a floating log and a fish swam beneath it."
- For on recognition was positive, for beside negative

Not the semantics of the sentence but the described situation is represented

John's Sweatshirt

- Glenberg, Meyer and Lindem, 1987
- Setting: John was preparing for a marathon in August.
- Critical: After doing a few warm-up exercises, he took on/put off his sweatshirt and went jogging.
- Filler: He jogged half-way around the lake without too much difficulty.
- Filler: Further along the route, however, John's muscles began to ache.
- Question: Was the marathon to be held in summer?
- Verification task: Does probe word (e.g. "sweatshirt") appear in the text?

Distance Effect



- if sbjs construct a mental model of the text, the associated texts should produce faster verification, esp. as the delay grows and the character moves further in distance from the initial location

Mental Models in Thinking

- Philip Johnson-Laird
- A reader creates a mental model of the text being read, which simulates the 'world' being described, as the reader understands/interprets it.
- ambiguous passages of text can lead to more than one competing mental model,
- However, passages of text that unambiguously produce a single mental model are easier to comprehend.

Reasoning with Mental Models

Syllogistic reasoning

- Construction of an integrated internal representation of the premisses
- "Read out" the conclusion

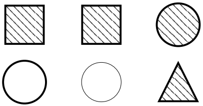
Influence factors on performance (error rates, performance time)

- Number of possible models
- Sequence of presentation of premises

Reasoning Example

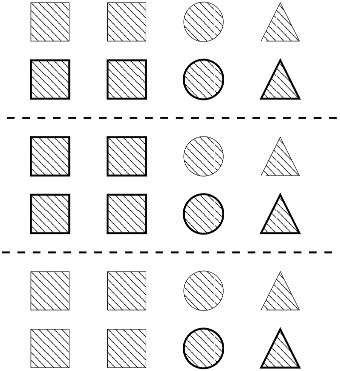
All squares are striped.
All striped objects have a bold margin.

All squares have a bold margin?



All squares are striped.
Some striped objects have a bold margin.

Some squares have a bold margin?



Seven Procedures for Reasoning Using Models

- Start a new model. Insert a new referent into the model according to a premise
- Update a model with a new relation to an existing referent
- Update a model with a new property or relation
- Join two separate models according to a relation between referents in them
- Verify whether a proposition is true or false in models
- Search for a counterexample to refute a proposition. If search fails, then the proposition follows validly from the previous propositions in the description
- Search for an example to make a proposition true. If search fails, then the proposition is inconsistent with the previous propositions.

(see Johnson-Laird and Yang in Sun, Computational Psychology)

Excursus: Reasoning

- Knowledge representation is the foundation for reasoning algorithms
- Reasoning can be deductive, inductive and abductive
- Reasoning can be logical, non-monotonic, probabilistic

Example for Deductive Reasoning: Wason Selection Task

(only 10% of subjects find the correct answer)

Deduction	
All humans are mortal.	(Axiom)
Socrates is human.	(Fact)
Conclusion:	
Socrates is mortal.	

Induction		Abduction	
Socrates is human.	(Background K.)	All humans are mortal.	(Theory)
Socrates is mortal.	(Observation(s))	Socrates is mortal.	(Obs.)
Generalization:		*Diagnosis:*	
All humans are mortal.		Socrates is human.	

Deduction: from general to specific ⇒ **proven** correctness

Induction: from specific to general ⇒ (**unproven**) knowledge gain

Abduction: from effect to cause ⇒ (**unproven**) diagnosis

Related Concepts of Mental Models

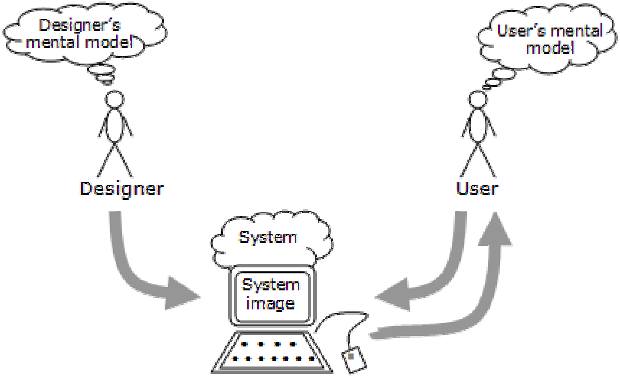
- In discourse: situation model (van Dijk and Kintsch, 1983)
- In naive physics (Forbus and Gentner,)
  see also qualitative reasoning in AI
- In HCI (D.Norman)

Mental Models in HCI

- Slower processing of ambiguous sentences: relevant to interaction design
- Interaction designers are interested in measures of learnability and ease of use (explanatory power of the MM theory)
- 'System causality conveyance' (Donald Norman, "The Design of Everyday Things", 1988)
- Description how a system is designed and implemented on the basis of the designer's mental model
- Similar to a reader of a passage of text, the user develops a mental model of how he/she thinks the system works through interaction of the system
- This model is used to reason about the system, to anticipate system behaviour and to explain why the system reacts as it does

Users' Mental Models

- The designer reifies (materialises) his mental model of a given design, e.g. a computer system, which becomes the only means of conveying his mental model to the user



Long-Term Memory

- Recognition is usually easier than recall
- Forgetting in LTM is usually attributed to bad retrieval cues and/or bad memory organization
- Sleep is thought to be improving consolidation of information
- Context-effects:
  Godden & Baddley (1975): learning of word lists under water or outside of water

Knowledge in LTM

- Declarative Knowledge (explicit, know what, verbalizable)
  - Semantic (concepts) $\hookrightarrow$ Knowledge Representation
  - Episodic (related to times and places)
- Non-Declarative
  - Procedural (implicit, know how)
  - Perceptual

Knowledge Representation

Symbolic Representations based on Logic

- Prolog, terminological logics, ontological logics
- Feature Models (Smith, Shoben, & Rips 1974)
- Semantic networks
- Frames/Schemes
- Semantic networks with inheritance
- Spreading activation networks (Collins & Loftus, 1975)
- Structural representations (mental models in working memory)

Similarity based Approaches

- Prototypes
- Exemplar Theory
- Multidimensional Scaling

Knowledge Representation

Statistical/Neural Approaches

- Artificial Neural Networks
- Bayesian Models

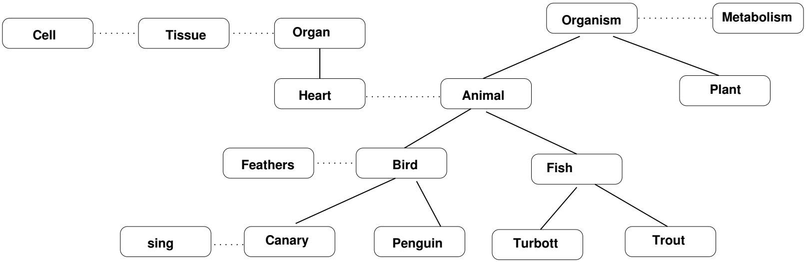
How is knowledge defined?

- Data: e.g. 42 (uninterpreted symbolic entities)
- Information: e.g., 42 degrees Celcius, 42 apples (interpreted entities)
- Knowledge: Mental representation of semantic content in relation to other concepts

Teachable Language Comprehender (TLC)

- Early network model of semantic memory
- Collins, A. M. & Quillian, M. R. (1969)
- Nodes represent concepts (like *Bird*)
- With each node is stored a set of properties (like *can fly*, *has wings*)
- With each node is stored a link to other nodes (like *canary* or *chicken*)
  ↪ **Hierarchical knowledge organisation**
- Properties are stored at the highest category level to which they apply
  ↪ inheritance – cognitive economy
- Can explain some empirical data, but not:
  - familiarity effect the typicality effect (*chicken is animal is answered faster than chicken is bird*)
    Update using weighted connections (Collins & Loftus, 1975)
  - Fast response to obvious negations (*a chicken is a car*)

Hierarchical Net



Semantic Networks in Prolog

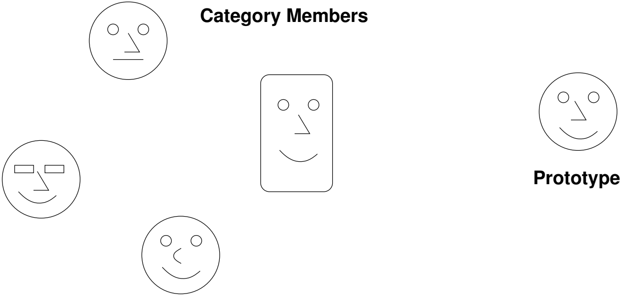
```
/* Eine Beispiel-Implementation eines semantischen      */
/*               Netzes in Standard-PROLOG              */
/* *****************************************************/

/* explizite Kanten im Netz  */
is_a(tier,lebewesen).
is_a(fisch,tier).
is_a(steinbutt,fisch).
is_a(herz,organ).
has_prop(tier,herz).
has_prop(organ,gewebe).
has_prop(gewebe,zellen).

/* ---------------------------------------------------- */
/* Ableitungsregeln im Netz                             */
/* A,B,C sind Konzepte; X,Y,Z sind Eigenschaften        */

isa(A,B) :- is_a(A,B).               /* R1: direkter Fall isa      */
isa(A,C) :- is_a(A,B), isa(B,C).     /* R2: Transitivität von isa  */

has(A,X) :- has_prop(A,X).           /* R3: direkter Fall has      */
has(X,Z) :- has_prop(X,Y), has(Y,Z). /* R4: Transitivität von has  */

has(A,X) :- has_prop(A,Y), isa(Y,X). /* R5: Verallg. von has bzgl.
                                            isa                    */
has(A,X) :- is_a(A,B), has(B,X).     /* R6: Vererb. von has bzgl. isa */
```

Prototypes

- Characteristic attributes (instead of defining attributes)
- e.g.: characteristic for a bird is building a nest, that it can fly, even if not all birds have these characteristics (penguin, ostric)
- Prototype theory (Medin, Rosch):
  - There is no attribute which must be shared by all members of a category, but there are characteristic attributes shared by large subsets of objects
  - "Family resemblance": cf. Wittgenstein "Spiel"
  - Prototype is an "average" object, having all characteristic features
  - Prototype itself typically has NO correspondence to a real object

Family Resemblance

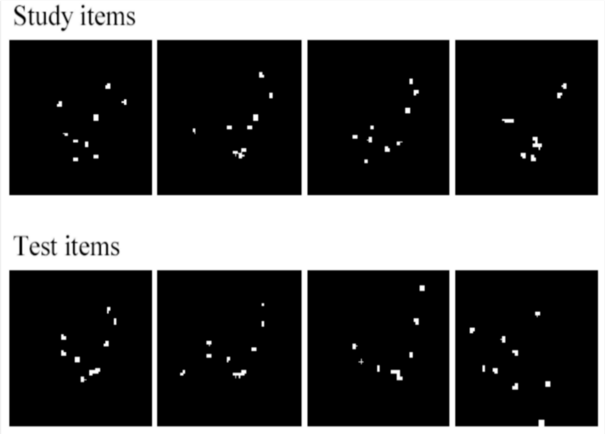
- Example (Medin et al. 2001)



Posner and Keele, 1968

- Study Items: High distortions of a prototype pattern
- Test Items: Prototype, low distortion, high distortion, random
- Result: Prototype is classified to belonging to the learned category!
- Interpretation: Similarity-based creation of prototypes as mean of the features of the exemplars

Dot Patterns

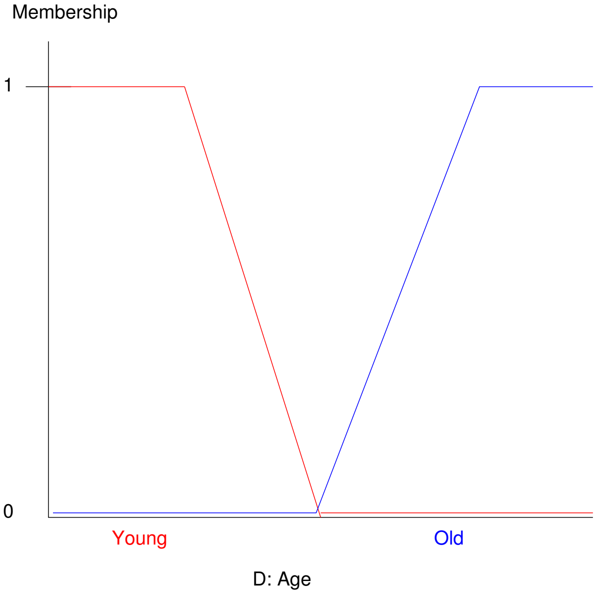


A Fuzzy-Set Model of Prototype Theory

L.A. Zadeh: Fuzzy Set Theory

- For a given domain D, elements x in a set A with different degrees of membership
- Characteristic function
  has to be predefined or determined empirically
  can be qualified with linguistic variables (e.g. *cold*, **very** *cold*)
- $c_A : D \to [0 \dots 1]$
- Compare to classical sets: $c_A : D \to \{0, 1\}$
- Fuzzy-Operators:
  Complement (not): $c_{nonA}(x) = 1 - c_A(x)$
  Intersection (and): $c_{A \cap B}(x) = min(c_A(x), c_B(x))$
  Union (or): $c_{A \cup B}(x) = max(c_A(x), c_B(x))$

Characteristic Function



Representation of Concepts

Osherson & Smith, 1981: Fuzzy Set Model of Prototype Theory

- Concept: $\langle A, d, p, c \rangle$
  with A: set of objects (concept)
  $d : A \times A \rightarrow R$ (distance metric)
  $p \in A$ (prototype)
  c (characteristic function)
- It holds that A is a metric space, i.e.
  $\forall x, y \in A$
  - $d(x, y) = 0$ iff $x = y$
  - $d(x, y) = d(y, x)$
  - $d(x, y) + d(y, z) \geq d(x, z)$
- It holds that objects closer to prototype are more characteristic, i.e.
  $\forall x, y \in A$
  $d(x, p) \leq d(y, p) \rightarrow c(y) \leq c(x)$

Conceptual Combination

Conjunctive Concepts

- e.g. *striped apple* (pet fish, red house, ...)
  $c_{stripedapple}(x) = min\{c_{striped}(x), c_{apple}(x)\}$
- it follows $c_{stripedapple}(x) \leq c_{apple(x)}$
- Psychologically, in the set of apples, a striped apple is less prototypical than a non striped apple

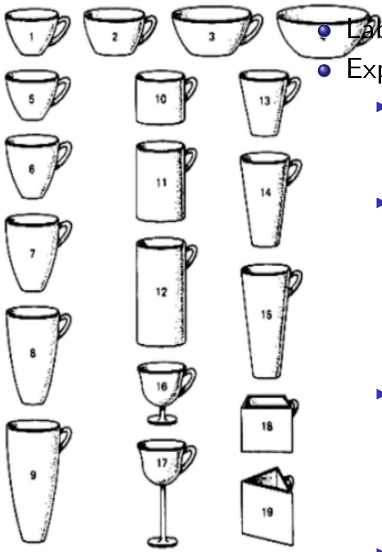
Excursus: Reasoning

Conjunction Fallacy

- Tversky & Kahneman: Human reasoners do often not follow rational rules
- e.g., rating a conjunctive concept as more probable than a component of this conjunction
- Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.
  Which is more probable?
  1. Linda is a bank teller.
  2. Linda is a bank teller and is active in the feminist movement.

Critique of Prototype Theory

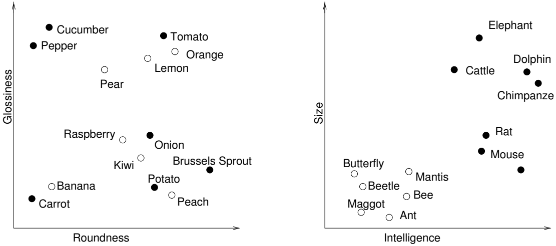
- If only a prototype is represented (exemplars forgotten), then
  - no information about variability (e.g. standard deviation of characteristic attributes)
  - no information about relative size of category
  - no consideration of attribute correlations (smaller birds typically can sing, larger birds not)
- Experiments show that humans use these kinds of information
- Context effects: e.g. typical beverage (office: coffee; construction site: beer)



- Labov
- Experiment: 19 pictures, tasks
  - Neutral: Pictures are shown in random order; subjects name them.
  - Same, but subject is asked to imagine that they saw someone with the object in his hand, stirring in sugar with a spoon, and drinking coffee or tea from it
  - they came to dinner at someone's house and saw this object sitting on the dinner table, filled with mashed potatoes/rice
  - standing on a shelf with cut flowers in it

Linear Separability

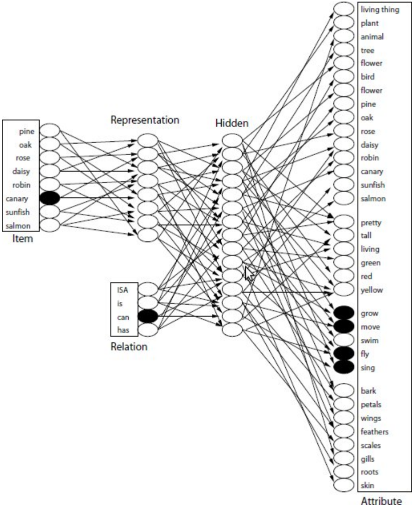
- Linear separability of categories: only for linear separable categories, prototypes can be defined which guarantee that the prototype is more similar to all exemplars belonging to the category than to exemplars belonging to another category
- Assumption: linear seperable categories are easier to learn (no conclusive empirical evidence)



Distributed Semantic Models

- Storing propositional knowledge in a PDP network
- PDP: special kind of artificial neural network approach, parallel distributed processing, made popular by McClelland & Rumelhart (1987)
- Hinton' distributed model (1981)
- Rumelhart model (1990): based on backpropagation

Rumelhart's Theory of Semantic Memory



Rumelhart's Theory of Semantic Memory

- Layered, feed-forward network
  Activate an item and a relation (value 1, all others 0)
  feed: $net_j = \sum_i a_i w_{ij}$
  bound between 0 and 1: $a = \frac{1}{1+e^{-net}}$
- Trained with backpropagation
  $err_p = \sum_i (a_{pi} - t_{pi})^2$
- Internal representation in the hidden unit

Summary

- Semantic memory:
  - stored in long term memory
  - in contrast to procedural memory
  - representation of generic declarative knowledge
- Computational models:
  - Hierarchical structure of symbolic representations
  - Prototypes
  - Distributed memory
- Access to semantic memory:
  - Retrieval
  - Activated knowledge: in working memory
  - e.g., Generating a mental model in working memory