Mental Models

As psychological representations of real, hypothetical, or imaginary situations, mental models were first postulated by the Scottish psychologist Kenneth Craik (1943), who wrote that the mind constructs "small-scale models" of reality to anticipate events, to reason, and to underlie EXPLANATION. The models are constructed in working memory as a result of perception, the comprehension of discourse, or imagination (see MARR 1982; Johnson-Laird 1983). A crucial feature is that their structure corresponds to the structure of what they represent. Mental models are accordingly akin to architects' models of buildings and to chemists' models of complex molecules.

The structure of a mental model contrasts with another sort of MENTAL REPRESENTATION. Consider the assertion

The triangle is on the right of the circle.

Its meaning can be encoded in the mind in a *propositional representation*, for example:

(right-of triangle circle)

The structure of this representation is syntactic, depending on the conventions governing the LANGUAGE OF THOUGHT: The predicate "right-of" precedes its subject "triangle" and its object "circle." In contrast, the situation described by the assertion can be represented in a mental model:

The structure of this representation is spatial: it is isomorphic to the actual spatial relation between the two objects. The model captures what is common to any situation where a triangle is on the right of a circle. Although it represents nothing about their distance apart or other such matters, the shape and size of the tokens can be revised to take into account subsequent information. Mental models appear to underlie visual IMAGERY. Unlike images, however, they can represent three dimensions (see MENTAL ROTATION), negation, and other abstract notions. The construction of models from propositional representations of discourse is part of the process of comprehension and of establishing that different expressions refer to the same entity. How this process occurs has been investigated in detail (e.g., Garnham and Oakhill 1996).

If mental models are the end result of perception and comprehension, they can underlie reasoning. Individuals use them to formulate conclusions, and test the strength of these conclusions by checking whether other models of the premises refute them (Johnson-Laird and Byrne 1991). This theory is an alternative to the view that DEDUCTIVE REASONING depends on formal rules of inference akin to those of a logical calculus. The distinction between the two theories parallels the one in LOGIC between proof-theoretic methods based on formal rules and model-theoretic methods based, say, on truth tables. Which psychological theory provides a better account of human reasoning is controversial, but mental models have a number of advantages. They provide a unified account of deductive, probabilistic, and modal reasoning. People infer that a conclusion is necessary -- it *must* be true -- if it holds in all of their models of

the premises; that it is probable -- it is *likely* to be true -- if it holds in most of their models of the premises; and that it is possible -- it *may* be true -- if it holds in at least one of their models of the premises. Thus an assertion such as

> There is a circle or there is a triangle, or both

yields three models, each of which corresponds to a *true* possibility, shown here on separate lines:

<div align="center">

O

Δ

O   Δ

</div>

The modal conclusion

> It is possible that there is both a circle and a triangle

follows from the assertion, because it is supported by the third model. Experiments show that the more models needed for an inference, the longer the inference takes and the more likely an error is to occur (Johnson-Laird and Byrne 1991). Models also have the advantage that they can serve as counterexamples to putative conclusions -- an advantage over formal rules of inference that researchers in artificial intelligence exploit in LOGICAL REASONING SYSTEMS (e.g., Halpern and Vardi 1991).

Mental models represent explicitly what is true, but not what is false (see the models of the disjunction above). An unexpected consequence of this principle is the existence of "illusory inferences" to which nearly everyone succumbs (Johnson-Laird and Savary 1996). Consider the following problem:

- Only one assertion is true about a particular hand of cards:
- There is a king in the hand or there is an ace, or both.
- There is a queen in the hand or there is an ace, or both.
- There is a jack in the hand or there is a ten, or both.
- Is it possible that there is an ace in the hand?

Nearly everyone responds "yes" (Johnson-Laird and Goldvarg 1997). Yet the response is a fallacy. If there were an ace in the hand, then two of the assertions would be true, contrary to the rubric that only one of them is true. The illusion arises because individuals' mental models represent what is true for each premise, but not what is false concomitantly for the other two premises. A variety of such illusions occur in all the main domains of reasoning. They can be reduced by making what is false more salient.

The term *mental model* is sometimes used to refer to the representation of a body of knowledge in long-term memory, which may have the same sort of structure as the models used in reasoning. Psychologists have investigated mental models of such physical systems as handheld calculators, the solar system, and the flow of electricity (Gentner and Stevens 1983).

They have studied how children develop such models (Halford 1993), how to design artifacts and computer systems for which it is easy to acquire models (Ehrlich 1996), and how models of one domain may serve as an ANALOGY for another domain. Researchers in artificial intelligence have similarly developed qualitative models of physical systems that make possible "commonsense" inferences (e.g., Kuipers 1994). To understand phenomena as a result either of short-term processes such as vision and inference or of long-term experience appears to depend on the construction of mental models. The embedding of one model within another may play a critical role in METAREPRESENTATION and CONSCIOUSNESS.

See also

- CAUSAL REASONING
- SCHEMATA

Additional links

- Mental Models
- Mental Models Website
- Mental models: a gentle guide for outsiders

-- Philip N. Johnson-Laird

References

Craik, K. (1943). *The Nature of Explanation.* Cambridge: Cambridge University Press.

Ehrlich, K. (1996). Applied mental models in human-computer interaction. In J. Oakhill and A. Garnham, Eds., *Mental Models in Cognitive Science.* Mahwah, NJ: Erlbaum.

Garnham, A., and J. V. Oakhill. (1996). The mental models theory of language comprehension. In B. K. Britton and A. C. Graesser, Eds., *Models of Understanding Text.* Hillsdale, NJ: Erlbaum, pp. 313-339.

Gentner, D., and A. L. Stevens, Eds. (1983). *Mental Models.* Hillsdale, NJ: Erlbaum.

Halford, G. S. (1993). *Children's Understanding: The Development of Mental Models.* Hillsdale, NJ: Erlbaum.

Halpern, J. Y., and M. Y. Vardi. (1991). Model checking vs. theorem proving: A manifesto. In J. A. Allen, R. Fikes, and E. Sandewall, Eds., *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference.* San Mateo, CA: Kaufmann, pp. 325-334.

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness.* Cambridge: Cambridge University Press; Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N., and R. M. J. Byrne. (1991). *Deduction.* Hillsdale, NJ: Erlbaum.

Johnson-Laird, P. N., and Y. Goldvarg. (1997). How to make the impossible seem possible. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society,* Stanford, CA. Hillsdale, NJ: Erlbaum, pp. 354-357.

Johnson-Laird, P. N., and F. Savary. (1996). Illusory inferences about probabilities. *Acta Psychologica* 93:69-90.

Kuipers, B. (1994). *Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge.* Cambridge, MA: MIT Press.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* San Francisco: Freeman.

Further Readings

Byrne, R. M. J. (1996). A model theory of imaginary thinking. In J. Oakhill and A. Garnham, Eds., *Mental Models in Cognitive Science.* Hove, England: Taylor and Francis. pp. 155-174.

Garnham, A. (1987). *Mental Models as Representations of Discourse and Text.* Chichester: Ellis Horwood.

Glasgow, J. I. (1993). Representation of spatial models for geographic information systems. In N. Pissinou, Ed., *Proceedings of the ACM Workshop on Advances in Geographic Information Systems.* Arlington, VA: Association for Computing Machinery, pp. 112-117.

Glenberg, A. M., M. Meyer, and K. Lindem. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language* 26:69-83.

Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:1084-1102.

Johnson-Laird, P. N. (1993). *Human and Machine Thinking.* Hillsdale, NJ: Erlbaum.

Legrenzi, P., V. Girotto, and P. N. Johnson-Laird. (1993). Focussing in reasoning and decision making. *Cognition* 49:37-66.

Moray, N. (1990). A lattice theory approach to the structure of mental models. *Philosophical Transactions of the Royal Society of London* B 327:577-583.

Polk, T. A., and A. Newell. (1995). Deduction as verbal reasoning. *Psychological Review* 102:533-566.

Rogers, Y., A. Rutherford, and P. A. Bibby, Eds. (1992). *Models in the Mind: Theory, Perspective and Application.* London: Academic Press.

Schaeken, W., P. N. Johnson-Laird, and G. d'Ydewalle. (1996). Mental models and temporal reasoning. *Cognition* 60:205-234.

Schwartz, D. (1996). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology* 30:154-219.

Stevenson, R. J. (1993). *Language, Thought and Representation.* New York: Wiley.