Cohen R (ed.) 1996 *Realism and Anti-realism in the Philosophy of Science*. Kluwer, Dordrecht, The Netherlands
Cornman J W 1975 *Perception, Common Sense, and Science*. Yale University Press, New Haven, CT
Devitt M 1984 *Realism and Truth*. 2nd edn. Princeton University Press, Princeton, NJ
Hacking I 1983 *Representing and Intervening*. Cambridge University Press, Cambridge, UK
Haldane J, Wright C (eds.) 1993 *Reality, Representation, and Projection*. Oxford University Press, Oxford, UK
Hooker C A 1987 *A Realistic Theory of Science*. SUNY Press, Albany, NY
Leplin J (ed.) 1984 *Scientific Realism*. University of California Press, Berkeley, CA
Maddy P 1990 *Realism in Mathematics*. Clarendon Press, Oxford, UK
McCormick P J 1996 *Starmaking. Realism, Anti-realism, and Irrealism*. MIT Press, Cambridge, MA
Niiniluoto I 1999 *Critical Scientific Realism*. Clarendon Press, Oxford, UK
Psillos S 1999 *Scientific Realism. How Science Tracks Truth*. Routledge, London
Putnam H 1981 *Reason, Truth and History*. Cambridge University Press, Cambridge, UK
Sayre-McCord G 1988 *Essays on Moral Realism*. Cornell University Press, Ithaca, NY
Sellars W 1963 *Science, Perception, and Reality*. Routledge, London
van Fraassen B 1980 *The Scientific Image*. Clarendon Press, Oxford, UK

U. Mäki

Reasoning with Mental Models

Reasoning is the mental process of drawing a conclusion from a set of premises. The premises may be statements, perceptions, beliefs, or items of general knowledge. The conclusion may be a statement or a thought that guides action. Reasoning is a central component of human intelligence, and without it, there would be no laws, civilization, or science (see *Scientific Reasoning and Discovery, Cognitive Psychology of*). Laws would have no application because individuals would be unable to make the following sort of inference:

Pat has a license or else Pat is disqualified from driving;
Pat does not have a license;
∴ Pat is disqualified from driving.

This inference is a valid deduction, that is, its conclusion must be true given that its premises are true. Logic is the science of valid deductions, but not a theory of how people reason. Such theories are the province of psychology, and psychologists have proposed a variety of theories (see *Logic and Cognition, Psychology of*). Human beings can reason about topics for which they have no general knowledge, and theories of this ability fall into two main categories. Some theorists postulate that reasoners rely unconsciously on formal rules of inference akin to those of a logical calculus (see *Natural Concepts, Psychology of*). The preceding inference depends on a rule of the form: A or else B, not A, therefore B. Other theorists propose that reasoners rely instead on their grasp of meaning and of principles akin to those for the semantics of a logical calculus. These theories rely on mental models, which are internal representations mirroring the structure of the external world. This article describes their role in reasoning and evidence corroborating it.

1. Mental Models

The idea that humans construct models of the external world goes back to the Scottish psychologist, Craik (see also *Mental Models, Psychology of*). He wrote:

> If the organism carries a 'small-scale model' of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present and the future, and in every way to react in a much fuller, safer, and more competent manner to the emergencies which face it (Craik 1943, Chap. 5).

Another antecedent is Wittgenstein's (1922) thesis that propositions represent reality in a similar way to pictures. What the modern theory of mental models adds to these programmatic proposals are three main assumptions:

(a) Each model represents a possibility. Thus, the assertion: either Pat has a license or else Pat is disqualified from driving calls for two models to represent the two possibilities (shown here in simplified form on separate horizontal lines):

License
Disqualified

where 'License' denotes a model of the possibility in which Pat has a license, and 'Disqualified' denotes a model of the possibility in which Pat is disqualified from driving.

(b) Models have a rich internal structure (not shown in the preceding diagram). Like an architect's plan, the parts of a model correspond to the parts of what it represents, and so the structure of a model corresponds to the structure of the world (Wittgenstein 1922, Propositions 2.13–2.17). Visual images are based on models, though many models are not visualizable (see *Imagery versus Propositional Reasoning*).

(c) The principle of truth: models normally represent what is true according to the premises, but not what is false. Hence, the preceding models of the disjunction represent only the possibilities that are true. Likewise, for true possibilities, models represent clauses in premises only when they are true in the possibility. For instance, the first model in the set above represents the possibility that Pat has a license,

but it does not represent explicitly that in this case it is false that Pat is disqualified from driving; likewise, the second model does not represent explicitly that in this case it is false that Pat has a license. Individuals make 'mental footnotes' about these matters, but normally they soon forget them. The principle of truth makes for parsimonious representations: reasoners do not have to bother with what is false.

Models are constructed as a result of perception, the comprehension of discourse, and imagination (see *Vision, Psychology of; Situation Model: Psychological*). They can represent real, hypothetical, or imaginary situations. They may also reside in long-term memory as a representation of knowledge (see *Mental Models, Psychology of*). They are accordingly a form of mental representation advocated by many psychologists as the way in which the mind represents reality, conceives alternative possibilities, and simulates the workings of the world. In principle, they could also underlie reasoning.

2. Deduction with Mental Models

The model theory of reasoning is due originally to the present author and his colleagues, notably the Irish psychologist Ruth Byrne (see Johnson-Laird and Byrne 1991, and for a recent alternative, Polk and Newell 1995). The theory postulates that models are based on the meaning of each premise, and that reasoners formulate a conclusion by describing something in the models that was not explicit in the premises. A conclusion that holds in all the models is necessary given the premises. A conclusion that holds in at least one model is possible given the premises. If models represent equiprobable possibilities—as reasoners often assume—the probability of a conclusion depends on the proportion of models in which it holds, as shown by the Italian psychologists Paolo Legrenzi, Maria Legrenzi, and Vittorio Girotto, and the French psychologist Jean-Paul Caverni (see Johnson-Laird et al. 1999). The model theory accordingly unifies logical reasoning, reasoning about possibilities, and probabilistic reasoning.

To illustrate the theory, consider again the inference:

Pat has a license or else Pat is disqualified from driving.

Pat does not have a license.

Therefore Pat is disqualified from driving.

The disjunctive premise calls for the models above. The second premise eliminates the model representing the possibility that Pat has a license. The remaining model is:

Disqualified

It yields the conclusion that Pat is disqualified from driving. This conclusion is a valid deduction because it holds in all the models—in this case, the single model—of the premises.

If individuals reason with models, then the more models that they have to construct—the more possibilities they have to envisage—the harder the task should be. Many experiments have corroborated this prediction (for a review including the studies described in this section, see Johnson-Laird 1999). For example, consider the following problem in spatial reasoning:

The cup is on the right of the knife.

The plate is on the left of knife.

The fork is in front of the plate.

The saucer is in front of the knife.

What is the relation between the fork and the saucer?

The theory predicts that reasoners should construct a two-dimensional model of the spatial layout of the items of the sort shown in the following plan:

plate knife cup
fork saucer

This model yields the conclusion: the fork is on the left of the saucer. In contrast, consider a problem in which the first premise is instead:

The knife is on the right of the cup

and the remaining premises are the same as above. This problem calls for two distinct models because the premises do not fix the relation between the plate and the cup:

plate	cup	knife	cup	plate	knife
fork		saucer		fork	saucer

The two models yield the same conclusion as before, but the problem should be harder because reasoners have to envisage two models. They do indeed draw a smaller percentage of correct conclusions to such multiple-model problems than to the one-model problems (Johnson-Laird and Byrne 1991).

Schaeken et al. have demonstrated similar differences between one-model and multiple-model problems about the temporal relations between events, using such premises as: John takes a shower before he drinks his coffee. Another group of Belgian psychologists (Vandierendonck et al.) have confirmed the difference in both temporal and spatial problems, and also shown that a task which preoccupies working memory for visual and spatial relations interferes with these inferences. Madruga et al. in Madrid have shown that inferences based on a conjunction, which calls for a single model, become harder when that premise is replaced by a disjunction, which calls for multiple models. Likewise, the Italian psychologists, Girotto et al. have shown that an inference based on a conditional premise of the form, If A then B, which calls for at least two models, becomes easier when a premise that eliminates one of these models is presented first rather than after the conditional premise. Two other Italian psychologists at Turin University, Bara and Bucciarelli, have shown that the difficulty of inferences known as syllogisms can also be predicted from the number of models that they require. Thus, the following syllogism is easy:

Some of the artists are beekeepers.

All of the beekeepers are chemists.
What follows?

Seven year-old children can draw the valid conclusion: some of the artists are chemists. In contrast, the next syllogism is hard even for adults:

None of the athletes is a banker.
All the bankers are chefs.
What follows?

Most reasoners construct a model in which the athletes are distinct from the chefs, and conclude that none of the athletes is a chef. To reach the correct conclusion, they need to realize that there could be chefs who are not bankers, and that each of the athletes could be such a chef. The only valid conclusion about chefs and athletes is accordingly that at least some of the chefs (i.e., those who are bankers) are not athletes. Few people reach this conclusion.

Another corroboration of the model theory comes from the study of children's understanding of assertions. Various experimenters, notably the French psychologists Barrouillet and Lecas, have demonstrated a trend in how children develop an understanding of conditional assertions of the form, If A then B. Young children treat them as though they were compatible with just a single model (in which both *A* and *B* occur), slightly older children treat them as compatible with two models (*A* and *B*, and not-*A* and not-*B*), and adolescents correctly treat them as compatible with three models (*A* and *B*, not-*A* and *B*, and not-*A* and not-*B*). Similarly, Sloutsky et al. at Ohio State University have observed that children tend to ignore the second clause of a disjunctive premise in order to ensure that the premise calls for just one model. This tendency to minimize the number of models has been corroborated by Ormerod et al. in their studies of how people paraphrase conditionals and other assertions.

A computer program implementing the model theory made a striking prediction. It showed that if reasoners abide by the principle of truth, then they should make systematic fallacies. These errors should arise because reasoners fail to represent what is false. Several recent studies have corroborated the occurrence of these 'illusory' inferences. Here is an example from an experiment carried out by Goldvarg and Johnson-Laird. Suppose that only one of the following assertions is true about a particular hand of cards:

There is a king in the hand or there is an ace in the hand, or both.

There is a queen in the hand or there is an ace in the hand, or both.

There is a jack in the hand or there is a ten in the hand, or both.

Is it possible that there is an ace in the hand?

In the experiment, 99 percent of the participants responded 'yes,' and the answer seems obvious. Yet, it is an illusion. It is impossible for an ace to be in the hand, because both of the first two assertions would then be true, contrary to the rubric that only one assertion is true. Other experiments have corroborated the occurrence of illusions in causal and conditional reasoning, and Yang, a Chinese psychologist, has shown that they occur with syllogistic premises.

The model theory and theories based on formal rules, which were mentioned earlier, run in parallel for many inferences, but illusory inferences are a case where the two accounts diverge. The model theory predicts the illusions, but the formal rule theories rely solely on valid principles of inference and so they are unable to explain them.

3. *Models, Induction, and Reasoning in Daily Life*

Many of the inferences that you make in daily life are not deductive, that is, even granted the truth of your premises, there is no guarantee that your conclusion is true. You leap to a conclusion that goes beyond the information in your premises. The conclusion rules out some additional possibilities over and above those that the premises rule out, and so it may be false. By definition, such inferences are known as inductions.

The model theory extends naturally to inductive inferences including the informal inductions of daily life (see *Informal Reasoning, Psychology of*). Induction consists in the addition of information to models, sometimes with the consequence of eliminating a model (of a possibility). Often the process is hardly distinct from the normal business of making sense of the world. For instance, you use a voltmeter to examine an electrical circuit and you discover that the following proposition is true:

The battery is dead or the voltmeter is faulty (or both).

You then find out that:

The voltmeter is faulty.

And so you infer: the battery is not dead. Your inference is an induction, because your conclusion could be false even though your premises are true. Your disjunctive premise has three models:

dead
faulty
dead faulty

where 'dead' is a model of the battery as dead and 'faulty' is a model of the voltmeter as faulty. Your discovery that the voltmeter is faulty is consistent with two of these models, and in one of them, the battery is dead (contrary to your conclusion). Hence, your conclusion goes beyond your premises. Yet, it is highly plausible. You may have based it on an unstated belief that the chances are remote of both a dead battery and a faulty voltmeter. This belief leads to you to discount a model so that only the single model yielding your conclusion survives.

Reasoning in real life almost always depends on background knowledge—a knowledge of the meaning of words, of the world, and of the situation at hand.

This knowledge enables you to leap to plausible conclusions. In logic, when a conclusion follows validly from premises, no subsequent information can invalidate it. Logic is monotonic in that additional premises lead merely to additional conclusions, and never to the withdrawal of an earlier conclusion. Logic means never having to be sorry about a valid conclusion. In daily life, however, you do withdraw a conclusion when it conflicts with subsequent evidence. Sometimes, you do so because the conclusion was based on an assumption that you made by default—an assumption warranted only if no evidence exists to the contrary, for example, Quakers are pacifists. If a President is a Quaker, then you may infer that he is a pacifist, but when you learn that he is war-mongering, you withdraw this conclusion. The model theory allows for the withdrawal of the consequences of default assumptions (Johnson-Laird and Byrne 1991).

Many other sorts of inference in daily life lead to conflicts with reality. Suppose you know, for example:

If I turn on the ignition then the engine will start.

You turn on the ignition, and so it follows validly:

The engine will start.

Unfortunately, nothing happens. There is a conflict between a valid consequence of your beliefs and the facts of the matter. Something has to give. At the very least, you have to withdraw your conclusion. Researchers in artificial intelligence have developed various systems of nonmonotonic reasoning to try to deal with such cases (see Brewka et al. 1997), but no one knows what mental processes underlie human nonmonotonic reasoning. At its heart, however, there appears to be the generation of diagnostic possibilities. You try to imagine a scenario—a mental model—that explains why the engine did not start.

4. *Conclusions*

The theory of mental models postulates that human beings reason by imagining the circumstances described in the premises. They envisage mental models of these possibilities, focusing on what is true at the expense of what is false. Unlike other theories of reasoning, this principle correctly predicts the occurrence of certain systematic fallacies. The evidence accordingly confirms that human reasoners do rely on mental models. They construct them from perception, memory, and imagination, manipulate them to make inferences, and base their actions on the results.

See also: Imagery versus Propositional Reasoning; Informal Reasoning, Psychology of; Logic and Cognition, Psychology of; Mental Models, Psychology of; Practical Reasoning: Philosophical Aspects; Problem Solving and Reasoning: Case-based; Problem Solving and Reasoning, Psychology of; Problem Solving: Deduction, Induction, and Analogical Reasoning; Scientific Reasoning and Discovery, Cognitive Psychology of

Bibliography

Brewka G, Dix J, Konolige K 1997 *Nonmonotonic Reasoning: An Overview*. CLSI Publications, Stanford University, Stanford, CA

Craik K J W 1943 *The Nature of Explanation*. Cambridge University Press, Cambridge, UK

Evans J St B T, Newstead S E, Byrne R M J 1993 *Human Reasoning: The Psychology of Deduction*. Lawrence Erlbaum, Hove, UK

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

Johnson-Laird P N 1999 Deductive reasoning. *Annual Review of Psychology* **50**: 109–35

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

Johnson-Laird P N, Legrenzi P, Girotto V, Legrenzi M, Caverni J-P 1999 Naive probability: A mental model theory of extensional reasoning. *Psychological Review* **106**: 62–88

Polk T A, Newell A 1995 Deduction as verbal reasoning. *Psychological Review* **102**: 533–66

Web page on mental models and the bibliography maintained by Byrne et al. (http://www.tcd.ie/Psychology/Ruth_Byrne/mental_models/)

Wittgenstein L 1922 *Tractatus Logico-Philosophicus*. Routledge & Kegan Paul, London

P. N. Johnson-Laird

Rechtsstaat (Rule of Law: German Perspective)

'*Rechtsstaat*' is a term of the German language and is not easily translated. '*Recht*' stands for 'law' and '*Staat*' for 'state'; in English and most other languages grammar or semantics stand in the way of combining both elements into a compound. The theme evoked by the term '*Rechtsstaat*' is often referred to as 'rule of law,' although the latter expression possesses a tradition of its own and carries connotations different from the German '*Rechtsstaat*.'

1. 'Rechtsstaat*': an Introduction*

Rechtsstaat concerns the age-old question of how to achieve order and freedom within a state. It connotes a balance between creating government authority powerful enough to keep peace both internally and

 ISBN: 0-08-043076-7