How Can Computers Get Common Sense?

Two of the founders of the field of artificial intelligence disagree on how to make a thinking machine

Despite all the marvelous things that computers can do today, they simply lack many of the qualities that are present in human intelligence—they don't even have common sense. And it is not at all clear how to program computers to give them common sense. Or, as experts in artificial intelligence put it, it is not clear how to represent common sense knowledge in a computer. "I think the AI [artificial intelligence] problem is one of the hardest science has ever undertaken," says Marvin Minsky of Massachusetts Institute of Technology, who is one of the founders of the field of AI.

There are, of course, computer programs that frequently are described as possessing artificial intelligence. Such programs can perform medical diagnoses, for example, or can predict where mineral deposits lie. These so-called expert systems are developed by computer scientists who glean a list of rules and procedures from human experts, such as doctors or mineral prospectors. And often the systems are quite useful. But they also are quite limited. "Much of the ordinary common sense ability to predict the consequences of actions requires going beyond the rules present in expert systems," says John McCarthy of Stanford University.

Theoreticians, however, have reached no consensus on how to solve the AI problem—on how to make true thinking machines. Instead, there are two opposing philosophical viewpoints and a flurry of research activity along these two directions. The different viewpoints were represented at a recent meeting* of the American Association for Artificial Intelligence by Minsky and by McCarthy, who also is a founder of the AI field and is an inventor of the term "artificial intelligence."

McCarthy believes that the way to solve the AI problem is to design computer programs to reason according to the well worked out languages of mathematical logic, whether or not that is actually the way people think. Minsky believes that a more fruitful approach is to try to get computers to imitate the way the human mind works which, he thinks, is almost certainly not with mathematical logic.

"I really think of myself as a psychologist," says Minsky, who reports that he gets his inspiration for attempting to represent knowledge in a computer by thinking about thinking, talking to psychologists and by going to playgrounds and questioning children who have not yet learned to conceal their thinking processes by couching their explanations in logical terms. From these investigations, he has become convinced that there is no single, simple way to explain human reasoning. "I think human intelligence is an accumulation of many different mechanisms and methods," he remarks. "I bet the human brain is a kludge."

So how do you put a jumble of poorly understood mechanisms and methods into a computer? Minsky believes that trying to represent the whole system with mathematical logic gets you into too many difficulties. "I've become convinced that the idea of 'fact' and the idea of 'truth' are no good. I think facts and truth are only good in mathematics and that's an artificial system. Logical systems work very well in mathematics, but that is a well-defined world. The only time when you can say something like, If a and b are integers, then a plus b always equals b plus a is in mathematics."

Minsky gives an example of the kind of difficulties that can occur if mathematical reasoning is applied to the real world. "Consider a fact like, 'Birds can fly.' If you think that common-sense reasoning is like logical reasoning then you believe there are general principles that state, 'If Joe is a bird and birds can fly then Joe can fly.' But we all know that there are exceptions. Suppose Joe is an ostrich or a penguin? Well, we can axiomatize and say if Joe is a bird and Joe is not an ostrich or a penguin, then Joe can fly. But suppose Joe is dead? Or suppose Joe has his feet set in concrete? The problem with logic is that once you deduce something you can't get rid of it. What I'm getting at is that there is a problem with exceptions. It is very hard to find things that are always true."

An alternative approach that Minsky developed is a system called frame (for framework) systems. It is a psychological approach. The idea is to put large collections of information into a computer—much more information than is ever needed to solve any particular problem—and then to define, in each particular situation, which details are optional and which are not. For example, a frame for "birds" might include feathers, wings, egg-laying, flying, and singing. In a biological context, flying and singing are optional; feathers, wings and egg-laying are not.

In frame systems, there is a collection of frame definitions which set the scene for common-sense reasoning. But the importance of the details in a frame can change if there is a change in purpose or goal. If you are walking in the woods, the importance of "flying" in your bird frame is substantial. If you are in Antarctica its importance is minimal. Or, in another type of example, you may have two different images of another person—one is as a business associate and the other is as a friend. If you cannot understand the person's behavior when you are viewing him as a business associate, you switch frames and try to understand his behavior by viewing him as a friend. In a sense, frame systems are like logic, but there is one important difference. Ordinarily, logic would not say which things are most important in which frame.

Minsky himself never actually sat down to program a computer to use frame systems, but one of his students did. Ira Goldstein, who is now at Hewlett-Packard in Palo Alto, developed a computer language which he calls FRL, for frame representation language, which he and his colleagues use in developing expert systems.

Originally, FRL represented only static objects. But Steven Rosenberg at Hewlett-Packard recently began extending the language so that it also represents the rules people employ for reasoning. With Rosenberg's extension of FRL, says Goldstein, "You can tie rules of reasoning to a particular domain of discourse. With FRL, we emphasize more the use of specific knowledge to guide reasoning. We place less emphasis on general reasoning mechanisms devoid of heuristic guidance."

"Minsky never liked logic," says McCarthy. "When difficulties with mathematical reasoning came up, he felt they killed off logic. Those of us who did like

*The National Conference on Artificial Intelligence, sponsored by the American Association for Artificial Intelligence, was held on 18 to 22 August at Carnegie-Mellon University and the University of Pittsburgh.

0036-8075/82/0924-1237$01.00/0

logic, and there weren't many, thought we should find a way of fixing the difficulties." Whether logical reasoning is really the way the brain works is beside the point, McCarthy says. "This is AI and so we don't care if it's psychologically real."

What McCarthy would like to do is to express common sense facts in the language of first order mathematical logic, meaning a language consisting only of variables and relation symbols such as "less than" or "mother of." "A proper axiomatization is one in which a proof exists for all conclusions that are ordinarily drawn from these facts," McCarthy remarks. "But what we know now about common sense is that that's asking for too much. You need another kind of reasoning—nonmonotonic reasoning."

Ordinary mathematical reasoning is monotonic in that if you have a set of premises and a set of conclusions, the set of conclusions is monotonic in the premises. If you add more facts, any conclusions you could draw without the additional facts are still valid with them. But common sense reasoning is often quite different from this mathematical logic. McCarthy explains, "If you know I have a car, you may conclude that you can ask me for a ride. If I tell you the car is in the shop, you may conclude you can't ask me for a ride. If I tell you it will be out of the shop in 2 hours, you may conclude you can ask me." As more premises are added, the conclusion keeps changing. "What's new is the possibility of formalizing nonmonotonic reasoning." That is, the possibility of using rules like those of mathematical logic to represent even nonmonotonic reasoning in a computer.

McCarthy calls his version of nonmonotonic reasoning circumscription. Unlike frame systems, circumscription is not yet being applied. "Circumscription is new and it is still changing continuously as a theoretical idea. There is still more theory to be done before it can be used in applications."

Circumscription is used to restrict a predicate as much as possible compatible with the facts that are being taken into account. After this has been done, the desired conclusions may follow by mathematical logic. For example, in the "Birds can fly" problem, McCarthy would use a predicate called "prevented from flying." In it, he would put any facts preventing flying that were being taken into account. These could include, for example, birds that are penguins or ostriches, as well as dead birds, or birds with their feet in concrete. Then the computer would reason, "If Joe is a bird and Joe is not a member of the set 'prevented from flying' then Joe can fly."

But is this circumscription a substitute for common sense? It certainly cannot take into account every contingency. It is easy to think of examples of nonflying birds, such as a bird with a broken wing, that a person with common sense would recognize as unable to fly but the computer would not.

"The conclusions we draw are risky, but that's inevitable," says McCarthy. "We can't invent all the hypotheses that might come to mind although we would like to take into account all the obvious things or, if a nonobvious fact becomes apparent, to take it into account. There is no reason to suppose we can make an omniscient computer program. We only want to make it as good as people."

Yet, McCarthy observes, "I admit that there are difficulties with circumscription. Suppose someone says, 'This bird is in a cage and is only prevented from flying on occasion.' That way lies madness. You can be forced to keep elaborating. The key thing about trying to formalize common sense is to avoid being forced to haggle."

Alternatives to logical reasoning also have their difficulties. Nils Nilsson of SRI International, who is president elect of the American Association for Artificial Intelligence, believes that "alternatives to logic all seem to be somewhat fuzzy and mushy. Some people think that's a virtue—they think that's what intelligence is all about. I don't see the evidence for that." In addition, says Nilsson, many of the people who try to develop systems that are alternatives to logic simply don't know much about logic. (Nilsson emphatically excludes Minsky from this group.) As a result, their alternative systems turn out to be mere subsets of logic. "Some of the things they invent are pale imitations of what logic can do," Nilsson remarks. "In some cases, there may be a little something extra, they may stick a little finger out. But the way to handle that is to extend logic. I think we should stand on the foundation that's been developed."

All efforts to solve the knowledge representation problem share two major obstacles, McCarthy explains. "The preliminary problem is to decide what knowledge to represent. The key thing that we have not got formulated is the facts of the common sense world." Then, even if researchers do manage to represent knowledge in computers, they still are faced with the problem of getting answers out of the computer in a reasonable time.

It is both Minsky's and McCarthy's opinion that the problem of common sense will need many new ideas to go further. But in the meantime, Minsky predicts, there will be immensely valuable spin-offs from attempts to solve the AI problem. This has been the pattern so far. Time sharing, word processing, the computer language LISP, symbolic manipulations by computers, all were developed by AI researchers in the course of their work on more basic problems. Minsky and McCarthy make an analogy with physics. As Minsky says, "It took 300 years from the time of Galileo to the discovery of quantum mechanics. You might ask, 'What took those guys so long?' " Yet all along there were important practical consequences of basic research in physics.

Of course, if the AI problem is solved, it will have enormous social consequences which Minsky, for one, worries about. "Do we need AI? There certainly is a dark side to any kind of advance and that's the question of whether societies can tolerate new systems. One of the things that AI threatens to do is to make work unnecessary. The dark question is, what will we do instead of work?"

Is it even possible to solve the AI problem—to design a computer that has common sense and intelligence? Minsky, McCarthy, and others in the field are convinced that the problem will be solved eventually. Asked why he holds this view, McCarthy answers, "The alternative is to say that there is an area of nature that is not reachable by science. And nothing in the history of science supports that hypothesis."

—GINA KOLATA

"I think the AI [artificial intelligence] problem is one of the hardest science has ever undertaken," says Marvin Minsky of Massachusetts Institute of Technology.