Basic Questions

Q. What is artificial intelligence?

A. It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.

Q. Yes, but what is intelligence?

A. Intelligence is the computational part of the ability to achieve goals in the world. Varying kinds and degrees of intelligence occur in people, many animals and some machines.

Q. Isn't there a solid definition of intelligence that doesn't depend on relating it to human intelligence?

A. Not yet. The problem is that we cannot yet characterize in general what kinds of computational procedures we want to call intelligent. We understand some of the mechanisms of intelligence and not others.

Q. Is intelligence a single thing so that one can ask a yes or no question ``Is this machine intelligent or not?''?

A. No. Intelligence involves mechanisms, and AI research has discovered how to make computers carry out some of them and not others. If doing a task requires only mechanisms that are well understood today, computer programs can give very impressive performances on these tasks. Such programs should be considered ``somewhat intelligent''.

Q. Isn't AI about simulating human intelligence?

A. Sometimes but not always or even usually. On the one hand, we can learn something about how to make machines solve problems by observing other people or just by observing our own methods. On the other hand, most work in AI involves studying the problems the world presents to intelligence rather than studying people or animals. AI researchers are free to use methods that are not observed in people or that involve much more computing than people can do.

Q. What about IQ? Do computer programs have IQs?

A. No. IQ is based on the rates at which intelligence develops in children. It is the ratio of the age at which a child normally makes a certain score to the child's age. The scale is extended to adults in a suitable way. IQ correlates well with various measures of success or failure in life, but making computers that can score high on IQ tests would be weakly correlated with their usefulness. For example, the ability of a child to repeat back a long sequence of digits correlates well with other intellectual abilities, perhaps because it measures how much information the child can compute with at once. However, ``digit span'' is trivial for even extremely limited computers.

However, some of the problems on IQ tests are useful challenges for AI.

Q. What about other comparisons between human and computer intelligence?

Arthur R. Jensen [Jen98], a leading researcher in human intelligence, suggests ``as a heuristic hypothesis'' that all normal humans have the same intellectual mechanisms and that differences in intelligence are related to ``quantitative biochemical and physiological conditions''. I see them as speed, short term memory, and the ability to form accurate and retrievable long term memories.

Whether or not Jensen is right about human intelligence, the situation in AI today is the reverse.

Computer programs have plenty of speed and memory but their abilities correspond to the intellectual mechanisms that program designers understand well enough to put in programs. Some abilities that children normally don't develop till they are teenagers may be in, and some abilities possessed by two year olds are still out. The matter is further complicated by the fact that the cognitive sciences still have not succeeded in determining exactly what the human abilities are. Very likely the organization of the intellectual mechanisms for AI can usefully be different from that in people.

Whenever people do better than computers on some task or computers use a lot of computation to do as well as people, this demonstrates that the program designers lack understanding of the intellectual mechanisms required to do the task efficiently.

Q. When did AI research start?

A. After WWII, a number of people independently started to work on intelligent machines. The English mathematician Alan Turing may have been the first. He gave a lecture on it in 1947. He also may have been the first to decide that AI was best researched by programming computers rather than by building machines. By the late 1950s, there were many researchers on AI, and most of them were basing their work on programming computers.

Q. Does AI aim to put the human mind into the computer?

A. Some researchers say they have that objective, but maybe they are using the phrase metaphorically. The human mind has a lot of peculiarities, and I'm not sure

anyone is serious about imitating all of them.

Q. What is the Turing test?

A. Alan Turing's 1950 article *Computing Machinery and Intelligence* [Tur50] discussed conditions for considering a machine to be intelligent. He argued that if the machine could successfully pretend to be human to a knowledgeable observer then you certainly should consider it intelligent. This test would satisfy most people but not all philosophers. The observer could interact with the machine and a human by teletype (to avoid requiring that the machine imitate the appearance or voice of the person), and the human would try to persuade the observer that it was human and the machine would try to fool the observer.

The Turing test is a one-sided test. A machine that passes the test should certainly be considered intelligent, but a machine could still be considered intelligent without knowing enough about humans to imitate a human.

Daniel Dennett's book *Brainchildren* [Den98] has an excellent discussion of the Turing test and the various partial Turing tests that have been implemented, i.e. with restrictions on the observer's knowledge of AI and the subject matter of questioning. It turns out that some people are easily led into believing that a rather dumb program is intelligent.

Q. Does AI aim at human-level intelligence?

A. Yes. The ultimate effort is to make computer programs that can solve problems and achieve goals in the world as well as humans. However, many people involved in particular research areas are much less ambitious.

Q. How far is AI from reaching human-level intelligence? When will it happen?

A. A few people think that human-level intelligence can be achieved by writing large numbers of programs of the kind people are now writing and assembling vast knowledge bases of facts in the languages now used for expressing knowledge.

However, most AI researchers believe that new fundamental ideas are required, and therefore it cannot be predicted when human-level intelligence will be achieved.

Q. Are computers the right kind of machine to be made intelligent?

A. Computers can be programmed to simulate any kind of machine.

Many researchers invented non-computer machines, hoping that they would be intelligent in different ways than the computer programs could be. However, they usually simulate their invented machines on a computer and come to doubt that the new machine is worth building. Because many billions of dollars that have been spent in making computers faster and faster, another kind of machine would have to be very fast to perform better than a program on a computer simulating the machine.

Q. Are computers fast enough to be intelligent?

A. Some people think much faster computers are required as well as new ideas. My own opinion is that the computers of 30 years ago were fast enough if only we knew how to program them. Of course, quite apart from the ambitions of AI researchers, computers will keep getting faster.

Q. What about parallel machines?

A. Machines with many processors are much faster than single processors can be. Parallelism itself presents no advantages, and parallel machines are somewhat awkward to program. When extreme speed is required, it is necessary to face this awkwardness.

Q. What about making a ``child machine'' that could improve by reading and by learning from experience?

A. This idea has been proposed many times, starting in the 1940s. Eventually, it will be made to work. However, AI programs haven't yet reached the level of being able to learn much of what a child learns from physical experience. Nor do present programs understand language well enough to learn much by reading.

Q. Might an AI system be able to bootstrap itself to higher and higher level intelligence by thinking about AI?

A. I think yes, but we aren't yet at a level of AI at which this process can begin.

Q. What about chess?

A. Alexander Kronrod, a Russian AI researcher, said ``Chess is the *Drosophila* of AI.'' He was making an analogy with geneticists' use of that fruit fly to study inheritance. Playing chess requires certain intellectual mechanisms and not others. Chess programs now play at grandmaster level, but they do it with limited intellectual mechanisms compared to those used by a human chess player, substituting large amounts of computation for understanding. Once we understand these mechanisms better, we can build human-level chess programs that do far less computation than do present programs.

Unfortunately, the competitive and commercial aspects of making computers play chess have taken precedence over using chess as a scientific domain. It is as if the geneticists after 1910 had organized fruit fly races and concentrated their efforts on breeding fruit flies that could win these races.

Q. What about *Go*?

A. The Chinese and Japanese game of *Go* is also a board game in which the players take turns moving. *Go* exposes the weakness of our present understanding of the intellectual mechanisms involved in human game playing. *Go* programs are very bad players, in spite of considerable effort (not as much as for chess). The problem seems to be that a position in *Go* has to be divided mentally into a collection of subpositions which are first analyzed separately followed by an analysis of their interaction. Humans use this in chess also, but chess programs consider the position as a whole.

Chess programs compensate for the lack of this intellectual mechanism by doing thousands or, in the case of Deep Blue, many millions of times as much computation.

Sooner or later, AI research will overcome this scandalous weakness.

Q. Don't some people say that AI is a bad idea?

A. The philosopher John Searle says that the idea of a non-biological machine being intelligent is incoherent. He proposes the [Chinese room argument](www-formal.stanford.edu/jmc/chinese.html) www-formal.stanford.edu/jmc/chinese.html The philosopher Hubert Dreyfus says that AI is impossible. The computer scientist Joseph Weizenbaum says the idea is obscene, anti-human and immoral. Various people have said that since artificial intelligence hasn't reached human level by now, it must be impossible. Still other people are disappointed that companies they invested in went bankrupt.

Q. Aren't computability theory and computational complexity the keys to AI? [Note to the layman and beginners in computer science: These are quite technical branches of mathematical logic and computer science, and the answer to the question has to be somewhat technical.]

A. No. These theories are relevant but don't address the fundamental problems of AI.

In the 1930s mathematical logicians, especially Kurt Gödel and Alan Turing, established that there did not exist algorithms that were guaranteed to solve all problems in certain important mathematical domains. Whether a sentence of first order logic is a theorem is one example, and whether a polynomial equations in several variables has integer solutions is another. Humans solve problems in these domains all the time, and this has been offered as an argument (usually with some decorations) that computers are intrinsically incapable of doing what people do. Roger Penrose claims this. However, people can't guarantee to solve *arbitrary* problems in these domains either. See my [Review of *The Emperor's New Mind* by Roger Penrose](). More essays and reviews defending AI research are in [[McC96a]()].

In the 1960s computer scientists, especially Steve Cook and Richard Karp developed the theory of NP-complete problem domains. Problems in these domains are solvable, but seem to take time exponential in the size of the problem. Which sentences of propositional calculus are satisfiable is a basic example of an NP-complete problem domain. Humans often solve problems in NP-complete domains in times much shorter than is guaranteed by the general algorithms, but can't solve them quickly in general.

What is important for AI is to have algorithms as capable as people at solving problems. The identification of subdomains for which good algorithms exist is important, but a lot of AI problem solvers are not associated with readily identified subdomains.

The theory of the difficulty of general classes of problems is called *computational complexity*. So far this theory hasn't interacted with AI as much as might have been hoped. Success in problem solving by humans and by AI programs seems to rely on properties of problems and problem solving methods that the neither the complexity

researchers nor the AI community have been able to identify precisely.

Algorithmic complexity theory as developed by Solomonoff, Kolmogorov and Chaitin (independently of one another) is also relevant. It defines the complexity of a symbolic object as the length of the shortest program that will generate it. Proving that a candidate program is the shortest or close to the shortest is an unsolvable problem, but representing objects by short programs that generate them should sometimes be illuminating even when you can't prove that the program is the shortest.

John McCarthy
2007-11-12