|  |
| --- |
| **Q 1. Go through Turing’s list of alleged “disabilities” of machines, identifying which have been achieved, which are achievable in principle by a program, and which are still problematic because they require conscious mental states.** |
| **Problematic**   1. Being kind: There are programs that are helpful, but to be kind means it needs some internal state. 2. Friendly: This also needs some internal state. 3. Tell right from wrong: There is a considerable research in applying AI to legal reasoning. There are tools that assist the lawyer in deciding a case and doing research. 4. Fall in Love: Very problematic. 5. Enjoy strawberries and cream: There are two aspects, first in the taste perception which has no work in AI. The second one is enjoying; this part clearly requires consciousness. 6. Be the subject of its own thought: The problematic word here is “Thought”. Many programs can process themselves, as when a compiler compiles itself. Perhaps closer to human self-examination is the case where a program has an imperfect representation of itself. 7. Have much diversity of behavior as man: No machine has achieved this.   **Achieved**   1. Have initiative: There are types of Intelligent autonomous agents that takes initiatives. 2. Resourceful: Means clever at finding ways of doing things. Many machines have achieved this. 3. Beautiful: Many industrial artifacts are evidence that machines can be beautiful. There are some programs that created art. 4. Have sense of humor: We know that no major effort to produce humor works. However, this seems to be achievable in principle. 5. Make mistakes: Intelligent software do make mistakes. 6. Make someone fall in love with it: This criterion is actually not too hard to achieve. Machines such as dolls and teddy bears been doing this to children. |
| **Q 2. Find and analyze an account in the popular media of one or more of the arguments to the effect that AI is impossible.** |
| * **Marketing**   AI lacks imaginative powers and cannot come up with new ideas. It cannot change a pattern to create something different. AI cannot suggest that we approach our customers with a marketing campaign focused on delivery or add a 100% money-back guarantee to our product. It can’t create what isn’t already a recognizable pattern.   * **Care**   Consider current chatbots such as Nadia, a life-like chatbot used by the Australian government to help people access the National Disability Insurance Scheme. Nadia is voiced by actress Cate Blanchett, and can interpret, with 85% accuracy, the visual expressions of the people it chats with online. However, it doesn’t understand (the way a human does) what the chats mean and can’t feel emotions that the human who is chatting may be conveying through their words. Nadia has zero emotional intelligence. |
| **Q 3. Attempt to write definitions of the terms “intelligence,” “thinking,” and “consciousness.” Suggest some possible objections to your definitions.** |
| **Consciousness** is an entity which is the subtlest of all. It is beyond body, mind, and intelligence. For example, suppose a patient is under a coma. He or she does not think or feel anything, and so his or her mind is not active. Hence it is the state of being aware of and responsive to one's surroundings.  **Intelligence** has been defined in many ways: higher level abilities (such as abstract reasoning, mental representation, problem solving, and decision making), the ability to learn, emotional knowledge, creativity, and adaptation to meet the demands of the environment effectively  **Thinking** means what thought is held in one’s mind about something. Thinking can be used with mind interchangeably when we are considered with human brain. It can also be said that the process of considering or reasoning about something. |
| **Q 4. Does a refutation of the Chinese room argument necessarily prove that appropriately programmed computers have mental states? Does an acceptance of the argument necessarily mean that computers cannot have mental states?** |
| No. Searle’s Chinese room thesis says that there are some cases where running a program that generates the right output for the Chinese room does not cause true understanding/ consciousness. The negation of this thesis is therefore that all programs with the right Philosophical Foundations output do cause true understanding/consciousness. |
| **Q 5. In the brain replacement argument, it is important to be able to restore the subject’s brain to normal, such that its external behavior is as it would have been if the operation had not taken place. Can the skeptic reasonably object that this would require updating those neurophysiological properties of the neurons relating to conscious experience, as distinct from those involved in the functional behavior of the neurons?** |
| The progress that has been made so far a limited class of restricted cognitive activities can be carried out on a computer, some much better than humans, most much worse than humans is very little evidence. If all cognitive activities can be explained in computational terms, then that would at least establish that cognition does not require the involvement of anything beyond physical processes. |
| **Q 6. Suppose that a Prolog program containing many clauses about the rules of British citizenship is compiled and run on an ordinary computer. Analyze the “brain states” of the computer under wide and narrow content.** |
| To some extent this question illustrates the slipperiness of many of the concepts used in philosophical discussions of AI. Here is our best guess as to how a philosopher would answer this question. Remember that “wide content” refers to meaning ascribed by an outside observer with access to both brain and world, while narrow content refers to the brain state only. |
| **Q 7.** **Alan Perlis [Perlis:1982] wrote, “A year spent in artificial intelligence is enough to make one believe in God”. He also wrote, in a letter to Philip Davis, that one of the central dreams of computer science is that “through the performance of computers and their programs we will remove all doubt that there is only a chemical distinction between the living and nonliving world.” To what extent does the progress made so far in artificial intelligence shed light on these issues? Suppose that at some future date, the AI endeavor has been completely successful; that is, we have built intelligent agents capable of carrying out any human cognitive task at human levels of ability. To what extent would that shed light on these issues?** |
| The progress that has been made so far—a limited class of restricted cognitive activities can be carried out on a computer, some much better than humans, most much worse than humans—is very little evidence. If all cognitive activities can be explained in computational terms, then that would at least establish that cognition does not require the involvement of anything beyond physical processes. Of course, it would still be possible that something of the kind is actually involved in human cognition, but this would certainly increase the burden of proof on those who claim that it is. |
| **Q 8. Compare the social impact of artificial intelligence in the last fifty years with the social impact of the introduction of electric appliances and the internal combustion engine in the fifty years between 1890 and 1940.** |
| The impact of AI has thus far been extremely small, by comparison. In fact, the social impact of all technological advances between 1958 and 2008 has been considerably smaller than the technological advances between 1890 and 1940. The common idea that we live in a world where technological change advances ever more rapidly is outdated. |
| **Q 9. I. J. Good claims that intelligence is the most important quality, and that building ultra-intelligent machines will change everything. A sentient cheetah counters that “Actually speed is more important; if we could build ultrafast machines, that would change everything,” and a sentient elephant claims “You’re both wrong; what we need is ultra-strong machines.” What do you think of these arguments?** |
| This question asks whether our obsession with intelligence merely reflects our view of ourselves as distinct due to our intelligence. One may respond in two ways. First, note that we already have ultrafast and ultra-strong machines (for example, aircraft and cranes) but they have not changed everything only those aspects of life for which raw speed and strength are important. Good’s argument is based on the view that intelligence is important in all aspects of life, since all aspects involve choosing how to act. Second, note that ultra-intelligent machines have the special property that they can easily create ultrafast and ultra-strong machines as needed, whereas the converse is not true |
| **Q 10. Analyze the potential threats from AI technology to society. What threats are most serious, and how might they be combated? How do they compare to the potential benefits?** |
| It is hard to give a definitive answer to this question, but it can provoke some interesting essays. Many of the threats are actually problems of computer technology or industrial society in general, with some components that can be magnified by AI—examples include loss of privacy to surveillance, and the concentration of power and wealth in the hands of the most powerful. As discussed in the text, the prospect of robots taking over the world does not appear to be a serious threat in the foreseeable future. |
| **Q 11. How do the potential threats from AI technology compare with those from other computer science technologies, and to bio-, nano-, and nuclear technologies?** |
| Biological and nuclear technologies provide much more immediate threats of weapons, yielded either by states or by small groups. Nano-technology threatens to produce rapidly reproducing threats, either as weapons or accidently, but the feasibility of this technology is still quite hypothetical. As discussed in the text and in the previous exercise, computer technology such as centralized databases, network-attached cameras, and GPS-guided weapons seem to pose a more serious portfolio of threats than AI technology, at least as of today. |
| **Q 12. Some critics object that AI is impossible, while others object that it is too possible and that ultra-intelligent machines pose a threat. Which of these objections do you think is more likely? Would it be a contradiction for someone to hold both positions?** |
| To decide if AI is impossible, we must first define it. In this book, we’ve chosen a definition that makes it easy to show it is possible in theory—for a given architecture, we just enumerate all programs and choose the best. In practice, this might still be infeasible, but recent history shows steady progress at a wide variety of tasks. Now if we define AI as the production of agents that act indistinguishably form (or at least as intelligently as) human beings on any task, then one would have to say that little progress has been made. |