

Read Turing's original paper on AI Turing:1950. In the paper, he discusses several objections to his proposed enterprise and his test for intelligence. Which objections still carry weight? Are his refutations valid? Can you think of new objections arising from developments since he wrote the paper? In the paper, he predicts that, by the year 2000, a computer will have a 30% chance of passing a five-minute Turing Test with an unskilled interrogator. What chance do you think a computer would have today? In another 50 years?

### **Completely irrelevant objections**

I think some of the objections are no longer relevant and not worthy of discussion. They are the extrasensory perception objection, the theological objection and the heads in the sand objection, the continuous nature of nervous systems.

### **Almost irrelevant objections**

The objection from consciousness feels brittle: what is it about machines that prevents them from being able to feel something? Is consciousness even necessary for intelligence, or maybe it is a consequence of intelligence. Most AI research (even general AI research) isn't concerned with consciousness and there's a reason for it. 2 The argument that our behavior is informal and machines cannot exhibit such behavior is also invalid. If there's anything I have learned being a software engineer, it's in fact hard to make programs behave predictably, even when rules are explicit (one could argue that I am just a terrible software engineer). And when the programs don't have explicit instructions governing its behavior but rather behavior is learned from data it's even more unpredictable.

The argument of machines can never do x. Some instances of this argument can be straight up ignored as they are not at all relevant to a machine being intelligent. The other instances actually have some arguments worth paying attention to. These x's are usually the tasks humans are good at machines not so much. It is funny how we never talk about the tasks machines are good at and humans not so much. We are not comfortable to see ourselves as stupid I believe.

Over the years, especially in the previous two decades, lot of instances of this argument have been invalidated. Defeating the go champion, advances in self-driving technology, human-level performance at atari games, expert medical diagnosis, superhuman image recognition, machine translation, image and text generation, all of which are not trivial and machines made significant progress.

Also, AI is not developed the same way human intelligence is developed and naturally certain problems are going to be easier and some harder. It is immature to say machines can never do x by handwavy arguments without any sort of proof.

It is more productive to rephrase such arguments as for example "machines currently cannot reason abstractly" and see what it takes for machines to do so. This is in fact how machines are able to do "x".

The objection raised by Lady Lovelace is that machines can never really do nothing new. And humans can? What is new anyway? What constitutes originality? Not all consequences of facts are "known" apriori and can lead to surprises.

## **Objections that are reasonable**

The mathematical objection presents the undecidability of some problems as a limit to the power of machines. However, without any proof, we claim that humans have no such limits. This argument still carries some weight as some problems that we might need to solve to solve intelligence can turn out to be undecidable. It's not obvious at this point whether undecidability results will be the reason we cannot engineer general intelligence. Even then, all is not lost.

## **Progress on Turing Test**

First of all, the Turing test is not a great way to measure intelligence and it is for a reason most research is dedicated to pass it. However, the performance of present day AI would largely depend on the details of how the test is going to be conducted. It would depend on factors such as the domain of conversation, expertise of the interrogator etc. I think the GPT-3 would be quite successful in passing the Turing test, perhaps 60% to 80% chance of passing a 5 minute test. In another 50 years, it's likely going to be an assignment in an undergrad CS course.