

Paper: Turing, A. (1950), Computing Machinery and Intelligence, Mind, LIX: 433–460

1. Summary of the paper:

- Question: “Can machines think?” (p. 433)
- Proposal: Replace question by a different question, viz. can digital computers pass a certain test? Turing test or imitation game; delivers operationalization of the original question; so machines are narrowed down to digital computers, which can emulate discrete state machines; the idea of thinking is replaced by the idea that the Turing test is passed.
- Turing test (standard interpretation): Can a human interrogator tell the difference between a human and computer interlocutor by have a written conversation of it (on average in a large set of trials)?
- Main claims by Turing:
 - The new question is an appropriate surrogate for the old one.
 - Machines will at some time pass the Turing test.
- Method:
 - Clarification of concepts:
 - Turing test
 - digital computer
 - discrete state machine
 - Handle objections against claims (mainly against test):
 - Theology: thinking requires a soul.
 - “Too dreadful to be true”
 - Mathematics: it can be shown that computers can’t correctly answer some questions (undecidability).
 - Consciousness: thinking requires consciousness, but computer don’t have it.
 - Computers can’t do certain things that humans can do and that are somehow part of thinking.
 - Computer can’t be creative/can’t surprise us (Lady Lovelace).
 - The human brain is not digital.
 - Human behavior cannot fully be described using rules.
 - Humans have extra-sensory perception, computers don’t.
 - positive case for second claim: learning machines may be able to pass the Turing test.

2. Discussion points from our discussion:

- It all depends on the more general question in the background: what is thinking? We need a definition of thinking to answer the question.
- Philosophers have given philosophical definitions, but it’s arguable that a definition of thinking should capture what ordinary people mean with “thinking”.

- However, many philosophers actually try to define “thinking” such that the definition captures what ordinary people mean with the term. This is called conceptual analysis.
- Still, as a concept tied to ordinary language, the term is to some extent vague.
- Turing proposed his test, which is very much focused on language.
- Turing’s arguments in favor of the test: In language, you can cover a lot of topics, domains of thinking; you can ask whether the system can do a calculation etc.; also, physical aspects are not relevant because the question is about thinking, not about appearance etc.
- Background: Behaviorism, a movement within psychology and philosophy. The idea was basically to focus on observable behavior rather than on mental states that are not directly accessible to others. There are different variants of behaviorism; for instance, behaviorism can be considered to be methodological advice, but also a claim about what terms such as “thinking” mean.¹
- There are basically two lines of critique against the Turing test:
- Objection 1: It doesn’t cover everything that matters for intelligence. For instance, pattern recognition on visual data (identify objects in pictures) is an achievement of intelligence, but it cannot be tested in the Turing test because the latter is restricted to text. However, it’s fairly easy to extend the Turing test for such dimensions.
- Objection 2: The Turing test focuses on outputs, outward behavior, but thinking is more; how it is done is decisive. This is not covered by the test.
- Common type of argument against the Turing test:
 - (Human) thinking involves X.
 - Machines cannot X.
 - Thus, machines cannot think.
 - Common ideas for X:
 - suffering; but this human, but not necessary for thinking
 - consciousness (phenomenal consciousness: having qualitative experiences; this is also discussed by Turing)
 - self-awareness (but might machines have this)
- A famous argument against the Turing test and machine intelligence has been given by Searle.² Idea basically: a human could use fixed rules operating on symbols of a foreign language to give answers in this language without understanding the language. The linguistic behavior of this human would look OK, but they wouldn’t understand the language by construction. This point is transferred to computers: They work on symbols but don’t have a clue on what the symbols are.
- One of our putting Searle’s point is to say that the Turing test doesn’t distinguish between thinking and a mere simulation of thinking.
- Objection: Humans who sleep and or are in coma don’t fulfill the conditions either. However, it’s debatable to what extent they really think. Also, Turing seems to think that the test is not necessary for thinking; rather, anyone passing the test would count as thinking.
- Objection: The interrogator could easily recognize a machine because the latter is faster in running certain computations. Turing’s answer: Yes, but this would mean that the machine has the wrong strategy; the best strategy is probably not to solve every task given by the interrogator, but to imitate human responses, which may just amount to “I don’t know”.
- Some version of the Turing test was at the basis of the Loebner prize.³ Some recent studies tried to subject ChatGPT to something like the Turing test.⁴

1 Graham, George, "Behaviorism", *The Stanford Encyclopedia of Philosophy* (Spring 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), <https://plato.stanford.edu/archives/spr2023/entries/behaviorism/>.

2 Searle, Richards J. (1980), Minds, brains, and programs, *Behavioral and Brain Sciences* 3 (3), 417–424.

3 https://en.wikipedia.org/wiki/Loebner_Prize.

4 Mei, Qiaozhu, et al. "A Turing test of whether AI chatbots are behaviorally similar to humans." *Proceedings of the National Academy of Sciences* 121.9 (2024): e2313925121.

Summary: A fundamental question regarding the status of AI is whether it really thinks. The question depends on what we mean by “thinking”. Although our conception of thinking may be a bit vague, it can be used to run arguments about whether machines can think. Turing suggested an operationalization of “thinking” leading to the Turing test, but there are many objections against the test.