



جامعة المستقبل  
AL MUSTAQBAL UNIVERSITY

# كلية العلوم قسم علوم الذكاء الاصطناعي

**LECTURE: (3)**

**Subject: Alan Turing Test**

**Level: First**

**Lecturer: Dr. Hussein Al-Zaidy**



## **Turing Test**

### **History**

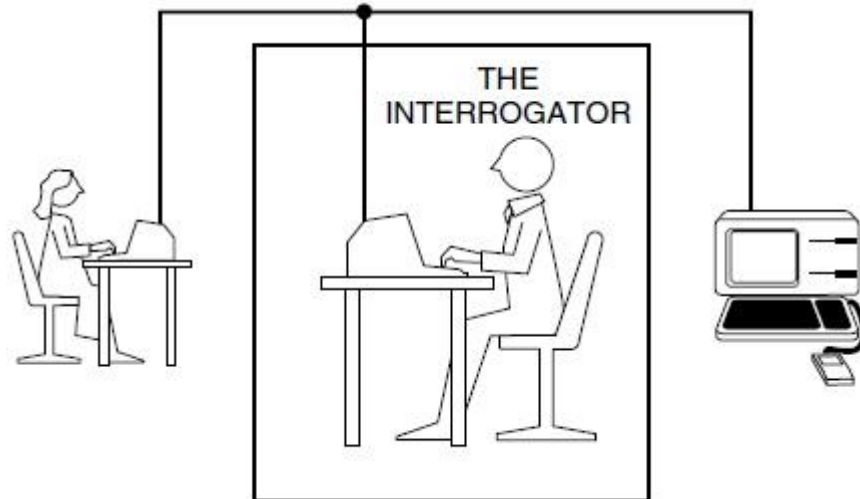
One of the earliest papers to address the question of machine intelligence specifically in relation to the modern digital computer was written in 1950 by the British mathematician Alan Turing. *Computing Machinery and Intelligence* (Turing 1950) remains timely in both its assessment of the arguments against the possibility of creating an intelligent computing machine and its answers to those arguments. Turing, known mainly for his contributions to the theory of computability, considered the question of whether or not a machine could actually be made to think. Noting that the fundamental ambiguities in the question itself (what is thinking? what is a machine?) precluded any rational answer, he proposed that the question of intelligence be replaced by a more clearly defined empirical test.

The Turing test measures the performance of an allegedly intelligent machine against that of a human being, arguably the best and only standard for intelligent behavior. The test, which Turing called the imitation game, places the machine and a human counterpart in rooms apart from a second human being, referred to as the interrogator (Figure 1). The interrogator is not able to see or speak directly to either of them, does not know which entity is actually the machine, and may communicate with them solely by use of a textual device such as a terminal. The interrogator is asked to distinguish the computer from the human being solely on the basis of their answers to questions asked over this device. If the interrogator cannot distinguish the machine from the human, then, Turing argues, the machine may be assumed to be intelligent.

Turing changes the concept of this game to include an AI, a human and a human questioner. The questioner's job is then to decide which is the AI and which is the human.



Since the formation of the test, many AI have been able to pass; one of the first is a program created by Joseph Weizenbaum called ELIZA.



### **Definition of Turing Test**

Different definitions of the turing test, but there is one more comprehensive which is:

"The Turing Test is a method of inquiry in artificial intelligence (AI) for determining whether or not a computer is capable of thinking like a human being. The test is named after Alan Turing, the founder of the Turing Test and an English computer scientist, cryptanalyst, mathematician and theoretical biologist".

### **Mechanisms of Turing Test**

Turing proposed that a computer can be said to possess artificial intelligence if it can mimic human responses under specific conditions. The original Turing Test requires three terminals, each of which is physically separated from the other two. One terminal is operated by a computer, while the other two are operated by humans.

During the test, one of the human's functions as the questioner, while the second human and the computer function as respondents. The questioner interrogates the



respondents within a specific subject area, using a specified format and context. After a preset length of time or number of questions, the questioner is then asked to decide which respondent was human and which was a computer.

The test is repeated many times. If the questioner makes the correct determination in half of the test runs or less, the computer is considered to have artificial intelligence because the questioner regards it as "just as human" as the human respondent.

### **The important features of Turing's test**

1. It attempts to give an objective notion of intelligence, i.e., the behavior of a known intelligent being in response to a particular set of questions. This provides a standard for determining intelligence that avoids the inevitable debates over its true" nature.
2. It prevents us from being sidetracked by such confusing and currently unanswerable questions as whether or not the computer uses the appropriate internal processes or whether or not the machine is actually conscious of its actions.
3. It eliminates any bias in favor of living organisms by forcing the interrogator to focus solely on the content of the answers to questions.

### **The impact of Turing's test**

Because of these advantages, the Turing test provides a basis for many of the schemes actually used to evaluate modern AI programs. A program that has potentially achieved intelligence in some area of expertise may be evaluated by comparing its performance on a given set of problems to that of a human expert. This evaluation technique is just a variation of the Turing test: a group of humans are asked to blindly compare the performance of a computer and a human being on a particular set of problems. As we will see, this methodology has become an essential tool in both the development and verification of modern expert systems.



## **How is the Turing Test used today?**

Although the variations of the Turing Test are often more applicable to our current understanding of AI, the original format of the test is still used to this day. For example, the Loebner Prize has been awarded annually since 1990 to the most human-like computer program as voted by a panel of judges. The competition follows the standard rules of the Turing Test. Critics of the award's relevance often downplay it as more about publicity than truly testing if machines can think.

At a competition organized by the University of Reading to mark the 60th anniversary of Turing's death in 2014, a Chatbot called Eugene Goostman that simulates a 13-year-old boy passed the Turing Test, in the eyes of some, when it fooled 33% of the judges. This so-called first pass has been met with much criticism from those who argue that there weren't enough judges, that other machines have performed better at the test in the past and that the test is invalid for only lasting five minutes.

In 2018, Google Duplex successfully made an appointment with a hairdresser over the phone in front of a crowd of 7,000. The receptionist was completely unaware that they weren't conversing with a real human. This is considered by some to be a modern-day Turing Test pass, despite not relying on the true format of the test as Alan Turing designed it.

GPT-3, a natural language processing model created by Open-AI, is thought by some to have the best chance of beating the test in its true form of any technology that we have today. But, even with its advanced text-generation abilities, many have criticized the machine because it can be tricked into answering nonsensical questions and therefore would struggle under the conditions of the Turing Test.

Despite much debate about the relevance of the Turing Test today and the validity of the competitions that are based around it, the test still stands as a philosophical starting



point for discussing and researching AI. As we continue to make advances in AI and better understand and map how the human brain functions, the Turing Test remains foundational for defining intelligence and is a baseline for the debate about what we should expect from technologies for them to be considered thinking machines.

### **Modern Approaches to the Turing Test**

Since the creation of the Turing test, more modern approaches have evolved in an attempt to better detect humans and machines. These variations of the Turing test are continually evolving to maintain relevance during technological advancements.

- The Reverse Turing Test aims to have a human trick a computer into having the computer believe it is not interrogating a human.
- The Total Turing Test incorporates perceptual abilities and the person being question's ability to manipulate objects.
- The Marcus Test has test subjects view media and respond to questions about the content consumed.
- The Lovelace Test 2.0 has test subjects create art and examines their ability to do so.
- The Minimum Intelligent Signal test asks test subjects only binary questions (i.e. only true/false or yes/no answers are allowed).

### **Has Any Machine Passed the Turing Test?**

As we mention early, in 2018, Google Duplex was introduced at the annual Google I/O Annual Developer Conference. The machine scheduled a hair salon appointment and interacted with a hair salon assistant via the phone as part of the conversation. Though some critics view the outcome differently, some believe Google Duplex passed the Turing test.



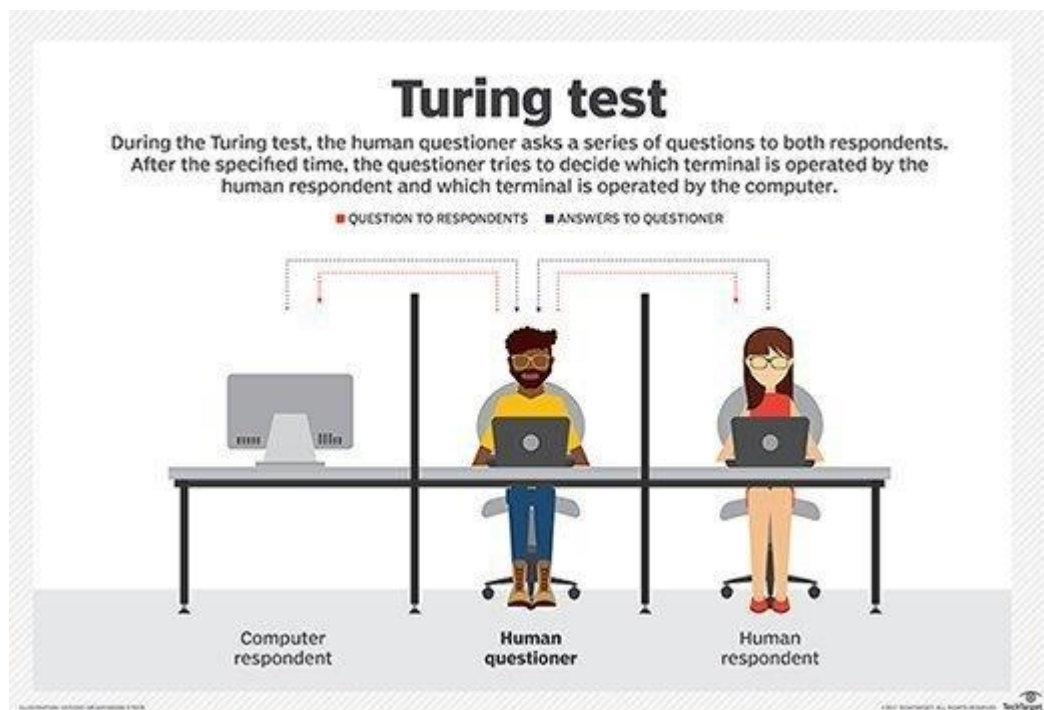
## Can a Human Fail the Turing Test?

Yes. Although a Turing test is based on knowledge and intelligence, it is also about evaluating how responses are given and whether the answers are interpreted to be sneaky.

For example, imagine being asked to provide the sum of 43,219 and 87,878. Whether or not you can provide the correct answer is only part of the exam; the Turing test evaluates how long it takes you provide an answer, any clarifying questions you ask in response, or whether you comprehend to add and not subtract the two figures. Based on any human's responses, it is possible to be mistaken for a computer (i.e. if you accidentally subtracted instead of added the figures, that may be incriminating evidence).

## The advantage and disadvantages of the Turing Test

There are many advantage and disadvantages of the using of the turing test.





The Turing Test has several potential benefits as a measure of machine intelligence. Some of the main advantages include:

- The test is relatively simple and easy to understand, making it accessible to a wide range of people.
- The test is based on the idea of human-like intelligence, which is a widely accepted and familiar concept.
- The test provides a clear and objective criterion for determining whether a machine is “intelligent” or not, which can help guide research and development in the field of artificial intelligence.
- The test can be used to compare the performance of different AI systems and track their progress over time.
- The test is a good way to evaluate the performance of AI in a way that is easy to replicate, compare and evaluate over time.

Additionally, the test provides a clear and objective criterion for determining whether a machine is “intelligent” or not which can help guide research and development in the field of artificial intelligence. It can be also a way to evaluate AI performance in an easy way to replicate, compare and evaluate over time.

On the other hand, there are several criticisms and limitations of the Turing Test as a measure of machine intelligence. Some of the main issues include:

- The test focuses solely on the ability to mimic human-like behavior and communication, rather than on the underlying intelligence or consciousness of the machine.
- The test is heavily dependent on the human evaluator’s subjective judgment, and may be influenced by factors such as the machine’s appearance or the human’s own biases.





- The test does not take into account the possibility that a machine could be intelligent in ways that are fundamentally different from human intelligence.
- The test does not consider the possibility of a machine deceiving the human evaluator, by providing pre-programmed or rehearsed responses rather than truly understanding the meaning of the questions.

Additionally, some argue that the test does not necessarily demonstrate true intelligence but rather the ability to simulate human-like intelligence and that it's not a true measure of machine's ability to think or being conscious.

The test is still widely discussed in the field of artificial intelligence. Overall, it can be said that the Turing test is a useful tool to measure certain aspects of AI but it should not be considered as the only tool or the ultimate measure of machine intelligence.