The Turing Test

Introduction

The Turing test is a central, long term goal for AI research –will we ever be able to build a computer that can sufficiently imitate a human to the point where a suspicious judge cannot tell the difference between human and machine? From its inception ithas followed a path similar to much of the AI research. Initially it looked to be difficult but possible (once hardware technology reached a certain point), only to reveal itself to be far more complicated than initially thought with progress slowing to the point that some wonder if it will ever be reached. Despite decades of research and great technological advances the Turing test still sets a goal that AI researchers strive toward while finding along the way how much further we are from realizing it. In 1950 English Mathematician Alan Turing published a paper entitled "Computing Machinery and Intelligence" which opened the doors to the field that would be called AI. This was years before the community adopted the term Artificial Intelligence as coined by John McCarthy. The paper itself began by posing the simple question, "Can machines think?".Turing then went on to propose a method for evaluating whether machines can think, which came to be known as the Turing test. The test, or "Imitation Game" as it was called in the paper, was put forth as a simple test that could be used to prove that machines could think. The Turing test takes a simple pragmatic approach, assuming that a computer that is indistinguishable from an intelligent human actually has shown that machines can think. The idea of such a long term, difficult problem was a key to defining the field of AI because it cuts to the heart of the matter –rather than solving a small problem it defines an end goal that can pull research down many paths. Without a vision of what AI could achieve, the field itself might never have formed or simply remained a branch of math or philosophy. The fact that the Turing test is still discussed and researchers attempt to produce software capable of passing it are indications that Alan Turing and the proposed test provided a strong and useful vision to the field of AI. It's relevance to this day seems to indicate that it will be a goal for the field for many years to come and a necessary marker in tracking the progress of the AI field as a whole. This section will explore the history of the Turing test, evaluate its validity, describe the current attempts at passing it and conclude with the possible future directions the Turing test solution may take.

Alan Turing

Alan Turing was an English mathematician who is often referred to asthe father of modern computer science[3]. Born in 1911, he showed great skill with mathematics and after graduating from college he published a paper "On Computable Numbers, with an Application to the Entscheidungsproblem" in which he proposed what would later be known as a Turing Machine –a computer capable of computing any computable function. The paper itself was built on ideas proposed by Kurt Godel that there are statements about computing numbers that are true, but that can't be proven*5+. Alan Turing worked on the problem in an effort to help define a system for identifying which statements could be proven. In the process he proposed the Turing Machine. The paper defines a "computing machine" with the ability to read and write symbols to a tape using those symbols to execute an algorithm [4]. This paper and the Turing machine provided that basis for the theory of computation.While Alan Turing focused primarily on mathematics and the theory of what would become computer science during and immediately after college, soon World War 2 came and he became interested in more practical matters. The use of cryptography by the Axis gave him reason to focus on building a machine capable of breaking ciphers. Before this potential use presented itself, Alan Turing likely hadn't been too concerned that the Turing machine he'd proposedin his earlier work was not feasible to build. In 1939 he was invited to join the Government Code and Cipher school as a cryptanalyst[5] and it became clear that he needed to build a machine capable of breaking codes like Enigma which was used by the Germans. He designed in a few weeks and received funding for the construction electromechanical machines called 'bombes' which would be used to break Enigma codes and read German messages by automating the processing of 126electrically linked Enigma scramblers. It wasn't the Turing machine, but the concepts of generating cyphertext from plaintext via a defined algorithm clearly fit with the Turing machine notion.After the war Turing returned to academia and became interested in the more philosophical problem of what it meant to be sentient, which lead him down the path to the Turing test.

Inception of the Turing Test

In 1950 Alan Turing was the Deputy Director of the computing laboratory at the University of Manchester. The paper which defined what would come to be known as the Turing test was published in a Philosophical journal called Mind. The paper itself was based on the idea of an 'Imitation Game'. If a computer could imitate the sentient behavior of a human would that not imply that the computer itself was sentient? Even though the description itself is fairly simple, the implications of building a machine capable of passing the test are far reaching. It would have to process natural language, be able to learn from the conversation and remember what had been said, communicate ideas back to the human and understand common notions, displaying what we call common sense.Similar to how he used the Turing Machine to more clearly formalize what could or could not be computed, Alan Turing felt the need topropose the Turing Test so that there was a clear definition of whether or not the responses given by a human were part of the computable space. In the paper he wanted to replace the question, 'Can machines think?' (which can have many possible answers and come down to a difference of opinion) with a version of the 'Imitation Game.' The original game upon which Turing's idea was based required a man, a woman and an interrogator. The goal was for the interrogator to identify which of the participants was a man and which was a woman. Since the interrogator would be able to identify the gender of the respondent by their voice (and maybe handwriting) the answers to the interrogator's questions would be type written or repeated by an intermediary. For theTuring Test, one of those two participants would be replaced by a machine and the goal of the interrogator would not be to identify the gender of the participants, but which is human and which is a machine.As described above, the Turing Test has a few key components that in effect define what Turing means when he wonders if machines can think. First the interrogator knows that there is one human and one machine. The test doesn't just require a computer to fool a human into thinking it is sentient; it asks the computer to fool a suspicious human. Second, physical nature isn't important –the goal is to not be able to tell the difference between man and machine when comparing the output of the machine and the true human. The communication medium is such that there are absolutely no hints beyond what can be expressed with written language. Also, the test doesn't include anything specific –no complex problem solving or requests to create art. As described, it seems a machine would pass the Turing test ifit were able make small talk with another human and understand the context of the conversation. For Turing, passing such a test was sufficient for him to believe that machines were capable of thinking.Beyond defining the game, the paper continues with an introduction to digital computers and how they can be used for arbitrary computation –harkening back to the description of the Turing machine. Taken with Godel's incompleteness theorem and Turing's formalization of what can and cannot be computed, the Turing test seems to strike at the simple question of whether that ability to appear sentient falls in to the realm of computable problems that a Turing machine can handle, or if it falls under the tiny subset of things that are true, but cannot be proven so. The test is simple, but the question is hugely significant and tied in to Turing's earlier work towards formalizing what can be computed.

Problems/Difficulties with the Turing Test

A large portion of Turing's original paper deals with addressing counter arguments concerning how the test he proposes may not be valid. In the introduction to that section he states that he believes there will be computers with enough storage capacity to make them capable of passing the Turing test "in about fifty years". The statement is interesting because it seems to imply that the AI software required to pass the Turing Test would be

7rather straightforward and that the limiting factor would only be memory. Perhaps this limitation was at the front of his mind because he was routinely running into problems that he could have solved if only there were enough storage available. The same type of reasoning is similar to what happens today when we believe that Moore's law will let us solve the hard problems. Beyond the storage limitations, he also raises other objections, including those based in theology (the god granted immortal soul is necessary for sentience), mathematical arguments based on Godel's work, the ability for humans to create original works and experience emotion, and others. One of the more interesting contradictions to the test is what he terms 'The Argument from Consciousness.' The argument goes that just imitating a human would not be enough because it doesn't invoke the full range of what it is that we consider to be human. Specifically, the Turing Test could be passed by a machine unable to do things such as write a poem or piece of music wrapped up as part of an emotional response. A machine passing the Turing test would not really have to experience or interpret art either. Turing argues that it is impossible to tell if the machine is feeling unless you are the machine, so there is no way to contradict the claim or to prove it. Using that method to dismiss the argument, he points out that the Turing test could include the machine convincing the interrogator that it is feeling something, even if there is truly no way to know that the emotions are actually being felt the way they would in a human. This would be similar to how humans communicate to convince each other of what they are feeling, though there is no guarantee that it is really true. Another interesting counter argument against the test that Turing describes is 'Lady Lovelace's Objection." She posited that because machines can only do what we tell them, they cannot originate anything, while it is clear that humans do originate new concepts and ideas all of the time. At the time this was written it may not have been possible to model the learning process, but much of the progress that has been made in teaching machines to learn and infer seems to have shown that this issue can be overcome. There have been specific implementations where voice or character recognition is reached by software training itself to recognize the variances in human writing or dialect. At least in these specific cases a machine can recognize something new so perhaps they will be able to in the general case as well. Overall the potential problems with the Turing test appear to fall in one of two categories: Does imitating a human actually prove intelligence or is it just a hard problem Is intelligence possible without passing the Turing test It seems fair to say that passing the Turing test is only a subset of the situation that humans have to contend with on a day to day basis. So it is possible that there are other key capabilities like experiencing emotions, having core beliefs or motivations, or problem solving that might be simulated in a computer but would not necessarily be the same as what humans do. The Turing test avoids these questions by judging the computer (and human) only on the text they output as part of the casual conversation that takes place during the test. So even if a computer could pass the Turing test, is that enough to say machines are 'intelligent' or that they can 'think', or does that just say that they can now pass the Turing test, and there is much more to understand before we do consider them intelligent. Beyond that, there are many humans that we'd consider sentient –young children for instance, that would probably do poorly in the Turing test because they haven't accumulated enough knowledge and experience in communication. We wouldn't apply the Turing test to them and say that they therefore are not capable of thought, which means that it might be possible for a computer to 'think' but still not pass the Turing test.