The Turing Test

Introduction to Cognitive Science

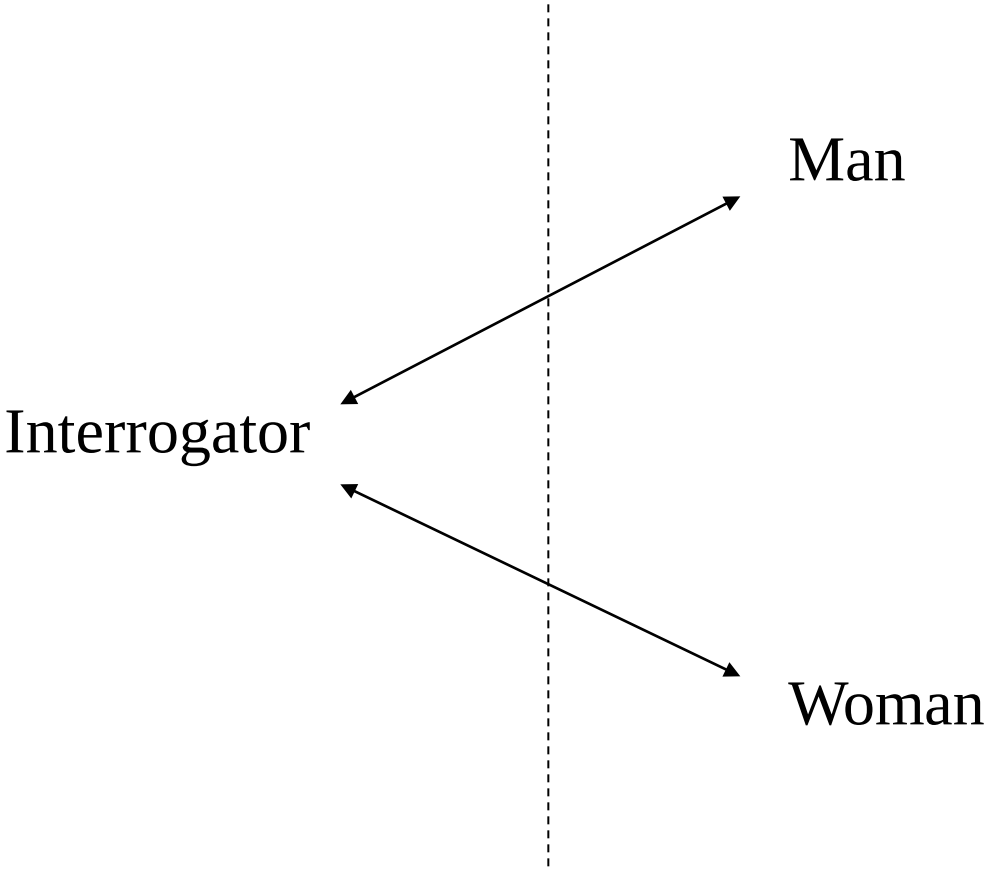
Computing Machinery and Intelligence (Turing, 1950)

I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous, If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

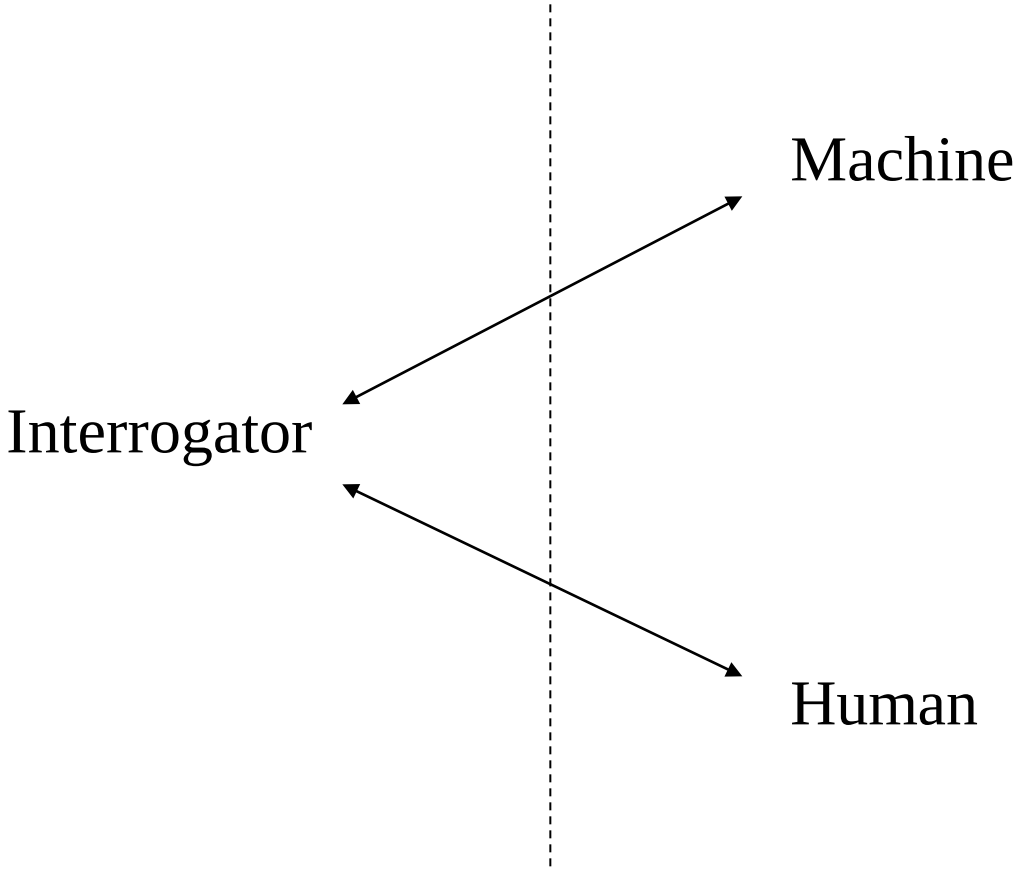
Computing Machinery and Intelligence (Turing, 1950)

The new form of the problem can be described in terms of a game which we call the 'imitation game." It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart front the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either "X is A and Y is B" or "X is B and Y is A.".

The Imitation Game



The New Imitation Game



“I believe that in about fifty years’ time it will be possible to programme computers, with a storage capacity of about 10^9, to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after 5 minutes of questioning”

-Alan Turing (1950)

The Turing Test

- Today the Imitation Game is usually referred to as the Turing Test.
- If a computer can play the game just as well as a human, then the computer is said to ‘pass’ the ‘test’, and should be declared intelligent.

Argument for Machine Intelligence Based on The Turing Test

- Anything that passes the Turing Test is intelligent
- Computers can pass the Turing Test
- Therefore, computers can be intelligent

Can Machines Think? The Behavioral Repertoire Argument

- Arguments for the possibility of thinking machines (or intelligent computers) often take the following form:
  - An entity is intelligent if it displays certain behavioral repertoires X
  - Computers can be programmed to display those behavioral repertoires X
  - Therefore, computers can be intelligent
- Clearly, the ‘Turing Test Argument’ fits this scheme.

Objections to this Argument

- While this argument is deductively valid, some people doubt it is well-founded:
  - *"Hollow Shell" Objection* –

    Premise 1 is questionable: Just because something displays certain behavioral repertoires X doesn't mean that it is intelligent; maybe it just behaves as if

  - *"Behavioral Shortcoming" Objection* –

    Premise 2 is questionable: I doubt that you can program a computer to do X

Objection to Computationalism: Simulations

- A computer simulation of a hurricane is just that: a simulation. It isn't a real hurricane!
- Similarly, simulating what a brain is doing is just that: a simulation of a brain, and not a real brain.

Response to the Simulation Objection

- Well, there are two notions of ‘simulation’:
  - ‘Computation’ (using a model): a computer can simulate a hurricane in that we are able to use a computer to model a hurricane and use that to compute certain things about hurricanes. However, *there is no mapping between the states that that computer goes through and that the hurricane goes through*. Similarly, simulating the actions of a brain merely gives us a description of the brain’s functioning.
  - ‘Emulation’: However, simulations in the above sense have *nothing* to do with the claim of computationalism which is about computations that *do* have the same functional organization as the brain, i.e. that emulate the brain!

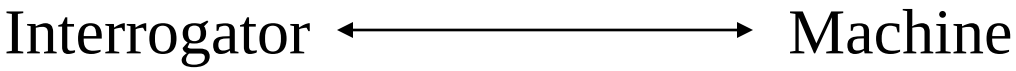
Some Initial Observations on the Turing Test

- The Turing Test attributes intelligence purely on verbal interactions. Is that ok?
- Well, physical characteristics (size, weight, agility, etc) don’t seem to be relevant as far as intelligence goes, so that seems right.
- However, shouldn’t we have to open up the computer program and see how it works to make this kind of determination?
- Then again, do we ever open up other human beings to determine whether they are intelligent?
- Hmm, maybe Turing has a point.

Why The Whole Set-Up?

- But if we're after a certain behavioral repertoire, why does the Turing Test have such a complicated set-up? Why did Turing 'pit' a machine against a human in some kind of 'imitation game'?
- That is, if Turing is trying to determine machine intelligence purely based on the interactions the interrogator is having with the computer's responses to certain questions, why not have the interrogator simply interact with a machine, see what it is or is not able to do, and determine whether or not the machine is intelligent based on those interactions? So why not:

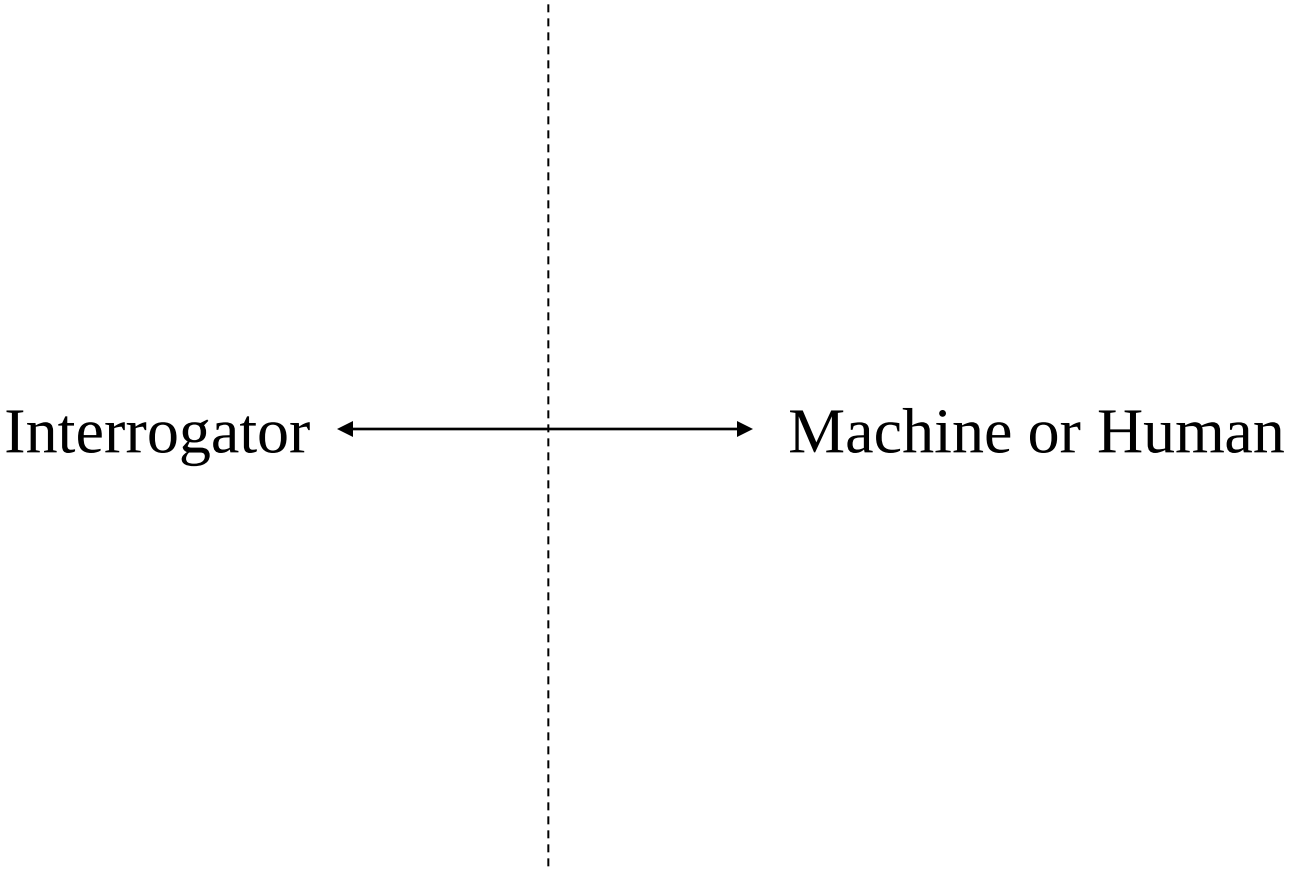
The Super-Simplified Turing Test!!



Answer: Bias

- The mere knowledge that we are dealing with a machine will bias our judgment as to whether that machine can think or not, as we may bring certain preconceptions about machines to the table.
- For example, knowing that we are dealing with a machine will most likely lead us to raise the bar for intelligence:
  - What, it can't write a sonnet? Aha! I knew it! It's not intelligent!
- By not knowing who or what is on the other end, such biases and raising-of-the-bar is eliminated in the Turing-Test.
- OK, but still, why not:

The Simplified Turing Test



Note: this is exactly how many commentators talk about the Turing Test!

Level the Playing Field

- Since we know we *might* be dealing with a machine, we *still* raise the bar for the entity on the other side being intelligent.
- (In fact, I bet that with this set-up probably a good number of humans would be declared to be machine!)
- Through his set-up of the test, Turing made sure that the bar for being intelligent wouldn't be raised any higher (or lower) for machines than we do for fellow humans.
- Thus, the Turing Test levels the playing field between humans and machines.

A Definition of Intelligence?

- Some commentators see the Turing Test as a definition of intelligence.
- And, many people have subsequently commented on the shortcomings of the Turing Test as a definition of intelligence:
  - This definition would amount to some kind of philosophical behaviorism. But, most of us think that while being intelligent *causes* the behavior, it does not *consist* in the behavior.
  - Also, this definition would be a real *sloppy* definition:
    - Who is the interrogator?
    - How long is the conversation?
    - What is the conversation about?
    - How does the interrogator decide?

Not a Definition

- Turing himself clearly did not intend to propose a definition of intelligence.
- In his paper Turing readily acknowledges that one could have intelligent beings not being able to pass the test simply by not having a human-like intellect:
  - “May not machines carry out something which ought to be described as thinking but which is very different from what a man does? This objection is a very strong one, but at least we can say that if, nevertheless, a machine can be constructed to play the imitation game satisfactorily, we need not be troubled by this objection”

A Sufficient Condition for Intelligence?

- Most commentators therefore interpret Turing's statement as saying that *if* a machine passes the Turing Test, *then* it is intelligent, i.e. that passing the Turing Test is a *sufficient* condition for intelligence (since intelligence is a necessary condition to pass it), but not a *necessary* one (and hence it is not a definition).
- In logic:
  - We have: $P \rightarrow I$
  - But not: $I \rightarrow P$

Same Sloppiness … And A Question

- But as a sufficient condition for being intelligent, the Turing Test suffers from some of the same problems as before:
  - such a criterion would still amount to a subjective judgment based on imprecisely defined behavioral criteria.
- In short, this seems to be a rather sloppy criterion!

Why would Turing (not exactly known for his sloppiness!) propose such a sloppy test?

Cheap Tricks? Eliza

- A psychotherapist program developed by Joseph Weizenbaum in 1966.
- Eliza used a number of simple strategies:
  - Keywords and pre-canned responses
    - “Perhaps I could learn to get along with my mother”

      -> “Can you tell me more about your family?”
  - Parroting
    - “My boyfriend made me come here”

      -> “Your boyfriend made you come here?”
  - Highly general questions
    - “In what way?”
    - “Can you give a specific example?”

Eliza and the Turing Test

- Many people conversing with Eliza had no idea that they weren't talking to a human.
- So did Eliza pass the Turing Test?
- (Or is it just easy being a psychotherapist?!)
- Eliza wasn't really tested in the format that Turing proposed.
- Still, it is interesting that humans were quick to attribute human-level intelligence to such a simple program.
- Maybe in a real Turing Test a relatively simple computer program can 'trick' the interrogator as well?

The Loebner Competition

- Modern day version of the Turing Test
- Multiple judges rank-order multiple humans and multiple computer programs from ‘most likely to be human’ to ‘least likely to be human’.
- Loebner has promised $100,000 for the first computer program to be ‘indistinguishable from a human’.
- Thus far, Loebner is still a rich man: occasionally a judge will rank a program above a human, but on the whole the judges systematically rank the humans above the computer programs.

An OK Test After All?

- Apparently it is quite difficult to pass the test!
  - When put to the real test, interrogators can see through superficial trickery
- So it seems we could say that if something does pass the test, then there is at least a *good chance* for it to be intelligent.
- In fact, if we are turning this into an inductive argument anyway, the sloppiness of the test isn't a huge concern either: we can now simply adjust our confidence in our claim in accordance to the nature of the conversation.
- So is this maybe what Turing was saying?

“Contrary Views”

- In his paper Turing goes over a list of “Contrary Views on the Main Question”:
- Machines:
  - Can’t be conscious
  - Can’t … (some specific ability, e.g. ‘be kind’, ‘use language properly’, ‘enjoy strawberries and icecream’, etc)
  - Can’t make mistakes
  - Can’t be creative
  - Can’t learn
  - Can’t do other than what they’re told (Lady Lovelace)

Machines can’t be conscious, have feelings, or have emotions

- Turing’s reply to this objection is that this response is most likely the result of a generalization of machines that we have encountered in our lives so far, all of which do indeed lack these qualities. However, it is not clear that in the future machines couldn’t have these qualities. So, without any further support for the truth of these claims, this objection really doesn’t work.
- Also, how do you know if a machines isn’t conscious? We don’t really know this for other humans either.

OK, Machines can do X, but they can't do Y (fill in anything for Y)

- Turing's reply to objections of this kind is that we should be careful not to require unreasonably much from the machine before we declare it to be intelligent. Many people are bad at playing chess or writing poetry, so if some machine can't do this, that doesn't automatically mean that the machine isn't intelligent. In fact, by using the Turing Test, Turing wanted to make sure that the bar wasn't raised any higher for machines than we do for fellow humans.
- Also, in saying that machines can't do Y, we may once again be generalizing from existing machines, rather than make any kind of argument for the in principle impossibility for machines to do Y.

Lady Lovelace Objection

- Probably the most common objection to machine intelligence: “Machines can only do what we tell them to do” (which is usually followed up by: “the program isn’t intelligent; the human programmer is!”)
- This takes several forms:
  - Machines can’t make mistakes
  - Machines can’t be creative
  - Machines can’t learn or adapt

Machines Can’t Make Mistakes

- Of course, this is a weird kind of objection to the possibility of machine intelligence, because what is so unintelligent about not making mistakes?
- Anyway, it is a proper objection to the claim that a machine could be able to pass the Turing Test, because supposedly a machine would always give itself away by its lack of mistakes (or by how inhumanly fast it is able to correctly solve math problems).

Response: Machines *Can* make Mistakes!

- However, machines sometimes do make mistakes (due to a bug in the program, say)
- In fact, it is easy enough to program a machine such that it does give the wrong answer to certain kinds of questions, and so that it does take a long time to give that answer.
- Of course, since we don't want a machine to make mistakes, we try to ensure that machines don't make mistakes. So, we rarely see machines making mistakes. But that doesn't mean that all machines are like that: we may once again be making a bad generalization based on the kinds of machines we see around us.

A Paradox: How can Machines make Mistakes?

- How can machines make mistakes given that they follow some deterministic routine, algorithm, or program?
- Turing: This paradox is easily resolved. It is indeed true that machines do exactly what their underlying mechanism, routine, or program dictates them to do. However, as a result of that routine, machines may end up getting the wrong answer, make the wrong decision, or do the wrong thing.

Levels or Perspectives

- Turing thus identified that machines can be looked at from two different levels or perspectives:
  - A ‘low-level operational’ level: looking at the machine ‘from the inside’; the machine following instructions (program) to manipulate symbols
  - A ‘high-level functional’ level: looking at the machine ‘as a whole’: the machine solving problems, answering questions, etc.

Machines can’t be Creative

- The point about machines not being able to make mistakes is often related to this one, the common ground being that *machines can only do what they are told to do*.
- The mistake in this objection is again that while this statement is true from the perspective of the underlying program, it is not clear that a machine couldn’t do anything new or creative when looked at from a higher level.
- Indeed, look at Deep Blue: Deep Blue beats every human in chess, but that would be impossible if Deep Blue couldn’t do any better than any of its programmers.

Machines can’t Learn

- Again, the underlying thinking here is that a machine can only do what it is told to do, and hence not do anything new and hence not learn.
- However, we know this claim to be false, because there are plenty of machines that do learn. E.g. there are machines that learn to play chess, learn to walk, learn to diagnose diseases, etc.
- Again, we only look at the machine from the underlying mechanical/programming point of view. Yes, the machine will follow some program and not deviate from it. However, as a result of doing so, it can learn.
- In general then, we can program a machine to make mistakes, do creative things, and learn. There is no contradiction there!

Analogies between Machines and Humans?

- One could say that humans are, like machines, subject to strict laws of nature: we can’t do anything other than what nature forces us to do.
- And, one could say that humans are ‘programmed’ by other humans (through education, etc.). Indeed, just because the programmer is intelligent, does that mean that the program is automatically not intelligent? How does that follow?

Another Question

- If Turing's point of his article was to propose a test or criteria for intelligence, then why are none of these objections about the validity of this test?
- In particular, given the nature of the test, one would expect a whole bunch of "Hollow Shell" objections, and as we saw, that is indeed what we got from the commentators (due to tricks or due to the subjective nature of the judgment, something can pass the test without being intelligent)
- But, at best, Turing's own list of objections seem to be "Behavioral Shortcoming" objections
- In fact, some of these objections don't even seem to really and directly address the behavioral repertoire that would be required to pass the test
- Indeed, almost all of Turing's paper seems to be a defense of the possibility of machine intelligence in and of itself.

- *So what was Turing's real point of the paper?*

Passing the Test

- Also, if Turing really would be more concerned with “Behavioral Shortcoming” Objections, then why is it that Turing hardly makes any effort to argue that machines can pass the test?
- In his paper, Turing merely lays out the principles of computation, and discusses the notion of universal computation, but Turing never directly addresses how this relates to passing the test.
- Presumably, Turing thinks that passing the test requires nothing more than some kind of information processing ability, which is exactly what computers do.

Yet Another Question

- But if that is true, then it seems that Turing could much more easily have argued as follows:
  - Intelligence requires nothing more than some kind of information processing ability
  - Computers can have this information processing ability
  - Therefore, computers can be intelligent
- Indeed, this is exactly how most proponents of AI make the argument today.

- *So why didn't Turing make this very argument? Why bring in the game at all?*

In Summary

- The “Contrary Views” make it clear that AI opponents think machines can’t do certain things, but Turing thinks they can.
- But: the Turing Test doesn’t seem to be able to shed any light on this issue: it just doesn’t seem to be at the center of this whole debate
- So:
- *If Turing really wanted to propose a test for machine intelligence, why not propose a test that much more directly and objectively tests certain abilities that both parties can agree on to be relevant to intelligence?*
- And:
- *If Turing wanted to defend the possibility of machine intelligence, why even bring up such a sloppy test at all?*
- Indeed:
- *What was the point of Turing’s paper?!?*

My Answer

- I propose that the convoluted set-up wasn't merely a practical consideration to eliminate bias in some strange game, but rather *the very point of his article*: I believe Turing wanted to reveal the closed-mindedness and unfair prejudice of people regarding the possibility of machine intelligence.

‘Imitation Game’ vs ‘Turing Test’

- In other words, I think it is likely that Turing never intended to propose any kind of test for machine intelligence (let alone propose a definition!).

- Interesting fact: In his original article Turing uses the word ‘pass’ or ‘passing’ 0 times, ‘test’ 4 times, and ‘game’ 37 times.

The Turing ‘Test’ as Harmful!

- In fact, I believe that seeing Turing’s contribution as laying out a test is harmful.
- The harm is that we have been thinking about the goal of AI in these terms, and that has been, and still is, detrimental to the field of AI.
- E.g. In “Essentials of Artificial Intelligence”, Ginsberg defines AI as “the enterprise of constructing a physical symbol system that can reliably pass the Turing Test”
- But trying to pass the test encourages building cheap tricks to convince the interrogator, which is exactly what we have seen with Eliza, Parry, and pretty much any entry in the Loebner competition.
- This kind of work has advanced the field of AI, and our understanding of intelligence … exactly zilch!
- So, I think we really should no longer refer to the Turing Test as the Turing ‘Test’!!

How to Read Turing's Paper

- So what did Turing really mean? Ultimately, this is an issue of history, and not an issue we, as cognitive scientists, need to be concerned about.
- Better questions to ask are: What, if anything, can we learn from Turing's paper? What would be a fruitful interpretation of his paper?
- Well, there are many interesting parts of the paper, especially in Turing's responses to the 'Contrary Views'.
- I also believe that seeing Turing's paper as laying out a genuine test is harmful, not helpful.
- Instead, I believe a fruitful reading of his paper is to see the Turing 'Test' as a statement about the use of the word 'intelligence'.

Artificial Flight and Artificial Intelligence

- Imagine going back 100 years when the Wright Brothers had their first flight.
- We can imagine people say: “Well, but that’s not real flight. There is no flapping of the wings!”
- But over time, we realized that it is, from the standpoint of using concepts that help us think and make sense of the world around us, a good idea to consider airplanes as really flying.
- Maybe the same is true for intelligence!

In Turing's Words

"The original question, 'Can machines think?', I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted."

-Alan Turing (1950)