

Activity 20

Conversations with computers—*The Turing test*

Age group Middle elementary and up.

Abilities assumed Answering general questions.

Time About 20 minutes.

Size of group Can be played with as few as three people, but is also suitable for the whole class.

Focus

Interviewing.

Reasoning.

Summary

This activity aims to stimulate discussion on the question of whether computers can exhibit “intelligence,” or are ever likely to do so in the future. Based on a pioneering computer scientist’s view of how one might recognize artificial intelligence if it ever appeared, it conveys something of what is currently feasible and how easy it is to be misled by carefully-selected demonstrations of “intelligence.”

Technical terms

Artificial intelligence; Turing test; natural language analysis; robot programs; story generation

Materials

We should have digital copies available for each student

A copy of the questions in the blackline master on page 225 that each child can see (either one for each pair of children, or a copy on an overhead projector transparency), and

one copy of the answers in the blackline master on page 226.

What to do

This activity takes the form of a game in which the children must try to distinguish between a human and a computer by asking questions and analyzing the answers. The game is played as follows.

There are four actors: we will call them Gina, George, Herb and Connie (the first letter of the names will help you remember their roles). The teacher coordinates proceedings. The rest of the class forms the audience. Gina and George are *go-betweens*, Herb and Connie will be answering questions. Herb will give a human's answers, while Connie is going to pretend to be a computer. The class's goal is to find out which of the two is pretending to be a computer and which is human. Gina and George are there to ensure fair play: they relay questions to Herb and Connie but don't let anyone else know which is which. Herb and Connie are in separate rooms from each other and from the audience.

What happens is this. Gina takes a question from the class to Herb, and George takes the same question to Connie (although the class doesn't know who is taking messages to whom). Gina and George return with the answers. The reason for having go-betweens is to ensure that the audience doesn't see how Herb and Connie answer the questions.

Before the class begins this activity, select people to play these roles and brief them on what they should do. Gina and George must take questions from the class to Herb and Connie respectively, and return their answers to the class. It is important that they don't identify who they are dealing with, for example, by saying "*She* said the answer is. . ." Herb must give his own short, accurate, and honest answers to the questions he is asked. Connie answers the questions by looking them up on a copy of the blackline master on page 226. Where the instructions are given in italics, Connie will need to work out an answer.

Gina and George should have pencil and paper, because some of the answers will be hard to remember.

1. Before playing the game, get the children's opinions on whether computers are intelligent, or if the children think that they might be one day. Ask for ideas on how you would decide whether a computer was intelligent.

I feel as if the George, Gina, Herb, and Connie will need to be us, especially if Gina and George need to readily be able to use a pencil and paper.

2. Introduce the children to the test for intelligence in which you try to tell the difference between a human and a computer by asking questions. The computer passes the test if the class can't tell the difference reliably. Explain that Gina and George will communicate their questions to two people, one of whom will give their own (human) answers, while the other will give answers that a computer might give. Their job is to work out who is giving the computer's answers.
3. Show them the **list of possible questions in the blackline master on page 225**. This can either be copied and handed out, or placed on an overhead projector.

Have them choose which question they would like to ask first. Once a question has been chosen, get them to explain why they think it will be a good question to distinguish the computer from the human. This reasoning is the most important part of the exercise, because it will force the children to think about what an intelligent person could answer that a computer could not.

Gina and George then relay the question, and return with an answer. The class should then discuss which answer is likely to be from a computer.

Repeat this for a few questions, preferably until the class is sure that they have discovered who is the computer. If they discover who is the computer quickly, the game can be continued by having Gina and George secretly toss a coin to determine if they will swap roles.

The answers that Connie is reading from are not unlike the ones that some “intelligent” computer programs can generate. Some of the answers are likely to give the computer away quickly. For example, no-one is likely to recite the square root of two to 20 decimal places, and most people (including, perhaps, the children in the class) would not be able to answer that question at all. Some questions will reveal the computer when their answers are combined. For example, the “Do you like...” answers sound plausible on their own, but when you encounter more than one it becomes apparent that a simple formula is being used to generate the answers from the questions. Some of the answers indicate that the question was misinterpreted, although the class might reason that the person could have made the mistake.

Many of the answers are very bland, but safe, and a follow-up question would probably reveal that the computer doesn't really understand the subject. Answering “I don't know” is reasonably safe for the computer, and might even make it seem more human—we would expect a child to answer “I don't know” to some of the questions too, such as the request for the square root of two. However, if a computer gives this answer too often, or for a very simple question, then again it would reveal its identity.

Since the goal of the computer is to make the questioners think that they are dealing with a person, some of the answers are deliberately misleading—such as the delayed and incorrect answers to the arithmetic problem. The questions and answers should provide plenty of fuel for discussion.

Again, this needs to be a digital copy for screen-readers

Question	Please write me a sonnet on the subject of the Forth Bridge.
Answer	Count me out on this one. I never could write poetry.
Question	Add 34957 to 70764.
Answer	<i>pause for about 30 seconds</i> ... 105621.
Question	Do you play chess?
Answer	Yes.
Question	My King is on the K1 square, and I have no other pieces. You have only your King on the K6 square and a Rook on the R1 square. Your move.
Answer	<i>after a pause of about 15 seconds</i> ... Rook to R8, checkmate.

Figure 20.1: Are the answers from a person or a computer?

Question	In the first line of the sonnet which reads “Shall I compare thee to a summer’s day,” would not “a spring day” do as well or better?
Answer	It wouldn’t scan.
Question	How about “a winter’s day”? That would scan all right.
Answer	Yes, but nobody wants to be compared to a winter’s day.
Question	Would you say Mr. Pickwick reminded you of Christmas?
Answer	In a way.
Question	Yet Christmas is a winter’s day, and I don’t think Mr. Pickwick would mind the comparison.
Answer	I don’t think you’re serious. By a winter’s day one means a typical winter’s day, rather than a special one like Christmas.

Figure 20.2: These answers are probably from a person!

Variations and extensions

The game can be played with as few as three people if Gina also takes the role of George and Connie. Gina takes the question to Herb, notes his answer, and also notes the answer from the blackline master on page 226. She returns the two answers, using the letters A and B to identify who each answer came from.

In order to consider whether a computer could emulate a human in the interrogation, consider with the class what knowledge would be needed to answer each of the questions on page 226. The children could suggest other questions that they would have liked to ask, and should discuss the kind of answers they might expect. This will require some imagination, since it is impossible to predict how the conversation might go. By way of illustration, Figures 20.1 and 20.2 show sample conversations. The former illustrates “factual” questions that a computer might be able to answer correctly, while the latter shows just how wide-ranging the discussion might become, and demonstrates the kind of broad knowledge that one might need to call upon.

There is a computer program called “Eliza” (or sometimes “Doctor”) that is widely available in several implementations in the public domain. It simulates a session with a psychotherapist, and can generate remarkably intelligent conversation using some simple rules. If you can get

hold of this program, have the children use it and evaluate how “intelligent” it really is. Some sample sessions with Eliza are discussed below (see Figures 20.3 and 20.4).

What’s it all about?

For centuries philosophers have argued about whether a machine could simulate human intelligence, and, conversely, whether the human brain is no more than a machine running a glorified computer program. This issue has sharply divided people. Some find the idea preposterous, insane, or even blasphemous, while others believe that artificial intelligence is inevitable and that eventually we will develop machines that are just as intelligent as us. (As countless science fiction authors have pointed out, if machines do eventually surpass our own intelligence they will themselves be able to construct even cleverer machines.) Artificial Intelligence (AI) researchers have been criticized for using their lofty goals as a means for attracting research funding from governments who seek to build autonomous war machines, while the researchers themselves decry the protests as a Luddite backlash and point to the manifest benefits to society if only there was a bit more intelligence around. A more balanced view is that artificial intelligence is neither preposterous nor inevitable: while no present computer programs exhibit “intelligence” in any broad sense, the question of whether they are capable of doing so is an experimental one that has not yet been answered either way.

The AI debate hinges on a definition of intelligence. Many definitions have been proposed and debated. An interesting approach to establishing intelligence was proposed in the late 1940s by Alan Turing, an eminent British mathematician, wartime counterspy and long-distance runner, as a kind of “thought experiment.” Turing’s approach was operational—rather than define intelligence, he described a situation in which a computer could demonstrate it. His scenario was similar to the activity described above, the essence being to have an interrogator interacting with both a person and a computer through a teletypewriter link (the very latest in 1940s technology!) If the interrogator could not reliably distinguish one from the other, the computer would have passed Turing’s test for intelligence. The use of a teletypewriter avoided the problem of the computer being given away by physical characteristics or tone of voice. One can imagine extending the exercise so that the machine had to imitate a person in looks, sound, touch, maybe even smell too—but these physical attributes seem hardly relevant to intelligence.

Turing’s original test was a little different from ours. He proposed, as a preliminary exercise, a scenario where a man and a woman were being interrogated, and the questioner had to determine their genders. The man’s goal was to convince the questioner that he was the woman, and the woman’s was to convince the questioner that she was herself. Then Turing imagined—for this was only proposed as a thought experiment—a computer being substituted for one of the parties to see if it could be just as successful at this “imitation game” as a person. We altered the setup for this classroom activity, because the kind of questions that children might ask to determine gender would probably not be appropriate, and besides, the exercise promotes sexual stereotyping—not to mention deception.

Imitating intelligence is a difficult job. If the roles were reversed and a person was trying to pass themselves off as a computer, they would certainly not be able to do so: they would be given away by their slow (and likely inaccurate) response to questions like “What is $123456 \times$

Person	Men are all alike.
Computer	In what ways?
Person	They're always bugging us about something or other.
Computer	Can you think of a specific example?
Person	Well, my boyfriend made me come here.
Computer	Your boyfriend made you come here.
Person	He says I'm depressed much of the time.
Computer	I'm sorry to hear that you are depressed.
Person	It's true. I'm unhappy.
Computer	Do you think coming here will help you not to be unhappy?
Person	I need some help, that much seems certain.
Computer	What would it mean to you if you got some help?
Person	Perhaps I could learn to get along with my mother.
Computer	Tell me more about your family.
	...
Computer	What else comes to mind when you think of your father?
Person	Bullies.
Computer	Does that have anything to do with the fact that your boyfriend made you come here?
	...

Figure 20.3: A conversation with a psychotherapist?

789012?"

However, it is surprisingly easy for a computer to acquire a thin veneer of conversational ability. A famous program developed in the early 60s, Eliza (mentioned earlier), simulated a non-directive psychotherapist in a dialog with the user. Figure 20.3 gives an example of a conversation with this system. Something approaching a natural conversation is achieved by devices such as

- generating canned responses (e.g. *Can you think of a specific example* in response to something beginning *They're always ...*; *I'm sorry to hear that ...* in response to *He says ...*);
- echoing the user's statements (e.g. *Your boyfriend made you come here* in response to *My boyfriend made me come here*);
- recognizing key words (*Tell me more about your family* in response to *... mother ...*);
- some stock phrases (*What would it mean to you if ...*);
- issues taken up from earlier parts of the conversation (*Does that have anything to do with the fact that ...*).

VP	Do you think that I can use the computer this morning?
Terminal	Why do you ask?
VP	I might be able to make some additional telecommunications sales.
Terminal	Why aren't you sure?
VP	My prospective customers have never seen a demonstration of the system.
Terminal	Does this mean very much to you?
VP	Of course it does.
Terminal	You seem very positive about that.
VP	This is ridiculous.
Terminal	What do you mean this is ridiculous?
VP	Is there any reason why I can't use the computer now?
Terminal	Tell me more about your prospective customers.
VP	Please dial me up on 491-1850
	<i>VP forgets to type period ... the computer doesn't answer ... VP becomes so infuriated that he calls the person at home.</i>
VP	Why are you being so snotty with me?
Person	What do you mean why am I being snotty with you?

Figure 20.4: A frustrating conversation

Of course, this does not mean that the Turing test has been passed, because the person is not asking questions with the aim of identifying whether the respondent is a machine or not: if they were, it would not be long before all was revealed.

A widely-quoted incident, which may be apocryphal but has now passed into the folklore of artificial intelligence, occurred when a computer sales vice-president allegedly came across a terminal that was normally connected directly to a particular person's house, but on this occasion happened to be running the "psychotherapist" program. Figure 20.4 shows what happened. Although the vice-president was evidently taken in, it is clear that the situation predisposed him to believe that he was talking to a person. If he had suspected otherwise, he would soon have found out!

Another system that appears to be able to hold intelligent conversations is a program called "SHRDLU", developed in the late 60s that accepted instructions to manipulate children's blocks on a table top. In fact, the blocks and the table were simulated and presented in the form of a picture on a computer screen, although it would not have been too difficult to make things work with a real robot (at least in principle). Figure 20.5 depicts the robot's micro-world. It can carry on surprisingly fluent conversations, as shown in the dialog of Figure 20.6. Amongst other things, this illustrates how it can obey orders, disambiguate instructions, work out the referents of pronouns, and understand new terms ("steeple").

However, the system is very fragile and the illusion it creates is easily shattered. Figure 20.7 shows a continuation of the conversation—which was not actually generated by the program, but certainly could have been. Although it was possible to define the new term *steeple* in Figure 20.6, only a very limited kind of term can be described because the robot has an extremely

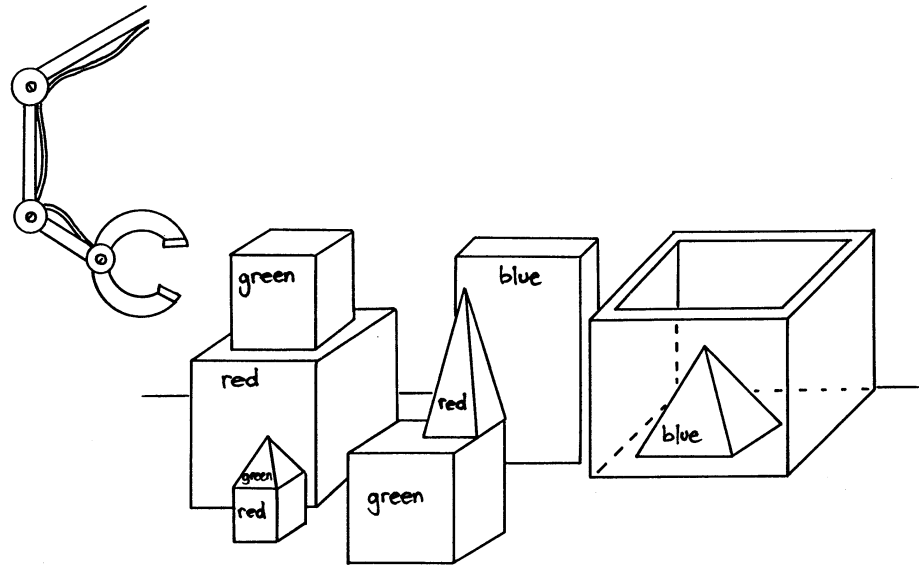


Figure 20.5: A robot's world of blocks

restricted “intelligence.” Ultimately, it has to admit that it doesn’t even understand the words it uses itself (“Sorry, I don’t know the word ‘sorry’ ”)! Now consider the alternate, completely imaginary, continuation of the conversation in Figure 20.8, which might have been generated by a child. Although tiny, this fragment demonstrates a wealth of world knowledge.

- The “robot” appreciates that a water pistol is less valuable if you already have one.
- Furthermore, it *expects the person to know that too* since it doesn’t bother to explain that this is a reason for declining the offer.
- Still, it sees that the person wants a block badly enough to part with something important.
- Also, it appreciates that—unlike water pistols—blocks are valuable in quantity.
- It tries to soften the person up by demeaning the frog.
- It implies that it is prepared to haggle.

Certainly this degree of artfulness is beyond today’s computers!

Figure 20.9 shows a story generated by a relatively recent computer program (late 80s). Although it is immediately given away as computer-generated by its sparsity and terseness, it is not hard to imagine how it could be dressed up by adding all sorts of extra detail. What is interesting is not the superficial aspects of the story but the plot that it embodies. While this is a long way from any human-generated plot, it does seem to capture some human elements of conflict.

Person	Pick up a big red block.
Robot	OK. <i>The system answers “OK” when it carries out a command. In order to pick up the red block, it had to clear it off by finding a space for the green one and moving the green one away.</i>
Person	Grasp the pyramid.
Robot	I don’t know which pyramid you mean. <i>There are three pyramids in the scene, so the program does nothing.</i>
Person	Find a block which is taller than the one you are holding and put it into the box. <i>The system has to figure out what is meant by “one” and “it”, as well as using deduction to decide which block it is holding.</i>
Robot	By “it”, I assume you mean the block which is taller than the one I am holding. <i>Printed out to keep us informed.</i>
Robot	OK. <i>The hand has to set down the red block to pick up the taller blue one.</i>
Person	What does the box contain?
Robot	The blue pyramid and the blue block.
Person	What is the pyramid supported by? <i>Earlier the system responded to “the pyramid” by saying “I don’t know which pyramid you mean.” Now, it knows that we must mean the one which was just mentioned.</i>
Robot	The box.
Person	Does a steeple . . .
Robot	Sorry, I don’t know the word “steeple”. <i>The system interrupts immediately when it sees a word it can’t figure out.</i>
Person	A “steeple” is a stack which contains two green cubes and a pyramid. <i>The quote marks indicate that we are defining a new word.</i>
Robot	I understand.
Person	Are there any steeples now?
Robot	No. <i>The new definition is used as a description to recognize steeples.</i>
Person	Build one.
Robot	OK. <i>The description given in the definition is enough to deduce a plan for building new objects as well as recognizing them. The word “one” means “a steeple”, and the plan is carried out.</i>

Figure 20.6: A conversation with a robot

Person	Trade you a water pistol for the big red block.
Robot	Sorry, I don't know the word "trade".
Person	A "trade" is a free exchange of property.
Robot	Sorry, I don't know the word "free".
Person	A "free" act is done willingly, without being forced.
Robot	Sorry, I don't know the word "act".
Person	"Acts" are what people do on purpose, and not by accident.
Robot	Sorry, I don't know the word "people".
Person	Sorry, I thought you were smarter than you are.
Robot	Sorry, I don't know the word "sorry".

Figure 20.7: (Imaginary) continuation of the conversation

Person	Trade you a water pistol for the big red block.
Robot?	I already have a water pistol, but I'll give you two blocks and a pyramid for your slimy old frog.

Figure 20.8: Another (imaginary) continuation of the conversation

Once upon a time there was an Arctic tern named Truman. Truman was homeless. Truman needed a nest. He flew to the shore. Truman looked for some twigs. Truman found no twigs. He flew to the tundra. He met a polar bear named Horace. Truman asked Horace where there were some twigs. Horace concealed the twigs. Horace told Truman there were some twigs on the iceberg. Truman flew to the iceberg. He looked for some twigs. He found no twigs. Horace looked for some meat. He found some meat. He ate Truman. Truman died.

Figure 20.9: A tale of conflict

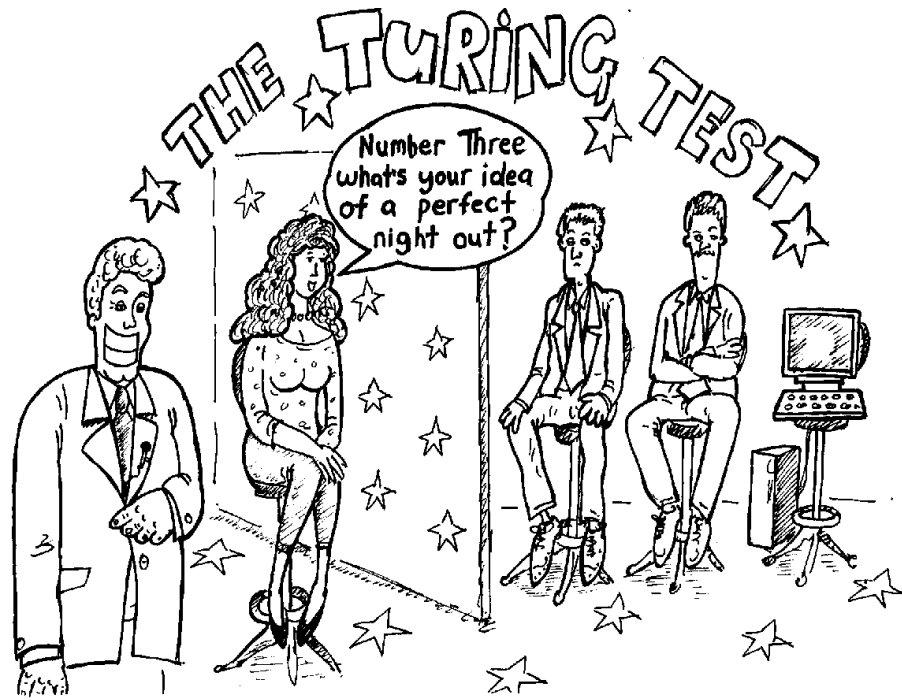
A competition using a restricted form of the Turing test was run in 1994, and some judges were fooled into thinking that a computer program was exhibiting intelligence. However, the judges were not allowed to use “trickery or guile,” and the topic of conversation was restricted. The restrictions tied the judges’ hands to the extent that some critics have argued that the test was meaningless. Like the activity above, restricting the paths that a conversation can take prevents the questioner from exploring areas that one would expect a natural conversation to take, and denies opportunities to demonstrate the spontaneity, creativity, and breadth of knowledge that are hallmarks of everyday conversation.

No artificial intelligence system has been created that comes anywhere near passing the full Turing test. Even if one did, many philosophers have argued that the test does not really measure what most people mean by intelligence. What it tests is behavioral equivalence: it is designed to determine whether a particular computer program exhibits the symptoms of intellect, which may not be the same thing as genuinely possessing intelligence. Can you be humanly intelligent without being aware, knowing yourself, being conscious, being capable of feeling self-consciousness, experiencing love, being . . . alive? The AI debate is likely to be with us for many more decades.

Further reading

Artificial intelligence: the very idea by the philosopher John Haugeland is an eminently readable book about the artificial intelligence debate, and is the source of some of the illustrations in this activity (in particular, Figures 20.7 and 20.8, and the discussion of them).

The original Turing test was described in an article called “Computing machinery and intelligence,” by Alan Turing, published in the philosophical journal *Mind* in 1950, and reprinted in the book *Computers and thought*, edited by Feigenbaum and Feldman. The article included Figures 20.1 and 20.2. The psychotherapist program is described in “ELIZA—A computer program for the study of natural language communication between man and machine,” by J. Weizenbaum, published in the computer magazine *Communications of the Association for Computing Machinery* in 1966. The blocks-world robot program is described in a PhD thesis by Terry Winograd which was published as a book entitled *Understanding natural language* (Academic Press, New York, 1972). The program that generated the story in Figure 20.9 is described in “A planning mechanism for generating story text,” by Tony Smith and Ian Witten, published in the *Proceedings of the 10th International Conference on Computing and the Humanities* in 1990. A competition involving a restricted form of the Turing Test is described by Stuart Shieber in *Communications of the Association for Computing Machinery*, June 1994.



1. What is the name of Bart Simpson's baby sister?
2. What do you think of Roald Dahl?
3. Are you a computer?
4. What is the next number in the sequence 3, 6, 9, 12, 15?
5. What do you think of nuclear weapons?
6. What is 2×78 ?
7. What is the square root of two?
8. Add 34957 to 70764.
9. Do you like school?
10. Do you like dancing?
11. What day is it today?
12. What time is it?
13. How many days are there in February in a leap year?
14. How many days are there in a week?
15. For which country is the flag a red circle on a white background?
16. Do you like to read books?
17. What food do you like to eat?

We need a digital copy of this to cooperate with screen readers.
This way the kids can peruse through and work as a group to ask questions.

Instructions: Choose questions from this list to ask the hidden human and "computer".

1. What is the name of Bart Simpson's baby sister?
I can't remember.
2. What do you think of Roald Dahl?
He writes funny books.
3. Are you a computer?
Are *you* a computer?
4. What is the next number in the sequence 3, 6, 9, 12, 15?
18.
5. What do you think of nuclear weapons?
Nuclear weapons are very dangerous and should not be used.
6. What is 2×78 ?
166 (*This is deliberately incorrect!*)
7. What is the square root of two?
1.41421356237309504878
8. Add 34957 to 70764.
Wait for about 20 seconds before giving the answer ... 105621.
9. Do you like school?
Yes, I like school.
10. Do you like dancing?
Yes, I like dancing.
11. What day is it today?
Give the correct day of the week.
12. What time is it?
Give the correct time.
13. How many days are there in February in a leap year?
2000 and 2004 are leap years.
14. How many days are there in a week?
Seven.
15. For which country is the flag a red circle on a white background?
I don't know.
16. Do you like to read books?
Yes, I like to read books.
17. What food do you like to eat?
I'm not hungry, thanks.

Instructions: *These are the answers to be used by the person pretending to be the computer.*