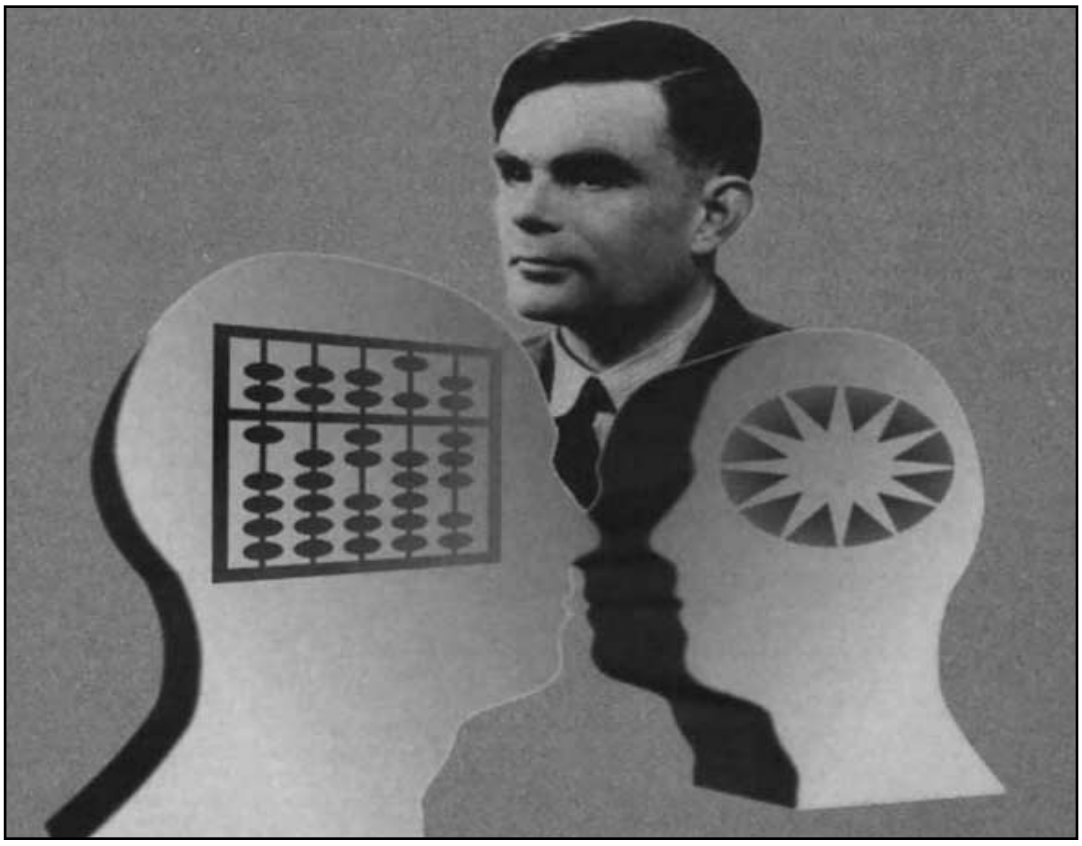
Can Machines Think?

Computers Try to Fool Humans at the First Annual Loebner Prize Competition Held at The Computer Museum, Boston



Can machines think?

Alan Turing's decades-old question still influences artificial intelligence because of the simple test he proposed in his article in Mind.

In this article, AI Magazine collects presentations about the first round of the classic Turing Test of machine intelligence, held November 8, 1991 at The Computer Museum, Boston.

Robert Epstein, Director Emeritus, Cambridge Center for Behavioral Studies, and an adjunct professor of psychology, Boston University, University of Massachusetts (Amherst), and University of California (San Diego) summarizes some of the difficult issues during the planning of this first real-time competition, and describes the event. He then speculates about the future of the competition and about its significance to the AI community.

Presented in tandem with Dr. Epstein's article is the actual transcript of session that won the Loebner Prize Competition—Joseph Weintraub's computer program PC Therapist.

0738-4602/92/$4.00 ©1992 AAAI

The Quest for the Thinking Computer

Robert Epstein, Boston University and the University of California, San Diego

In 1985 an old friend, Hugh Loebner, told me excitedly that the Turing Test should be made into an annual contest. We were ambling down a Manhattan street on our way to dinner, as I recall. Hugh was always full of ideas and always animated, but this idea seemed so important that I began to press him for details, and, ultimately, for money. Four years later, while serving as the director of the Cambridge Center for Behavioral Studies, an advanced studies institute in Massachusetts, I established the Loebner Prize Competition, the first serious effort to locate a machine that can pass the Turing Test. Hugh had come through with a pledge of $100,000 for the prize money, along with some additional funds from his company, Crown Industries, to help with expenses. The quest for the thinking computer had begun.

In this article, I'll summarize some of the difficult issues that were debated in nearly two years of planning that preceded the first real-time competition. I'll then describe that first event, which took place on November 8, 1991, at The Computer Museum in Boston and offer a summary of some of the data generated by that event. Finally, I'll speculate about the future of the competition—now an annual event, as Hugh envisioned—and about its significance to the AI community.

Planning

Planning for the event was supervised by a special committee, first chaired by I. Bernard Cohen, an eminent historian of science who had long been interested in the history of computing machines. Other members included myself, Daniel C. Dennett of Tufts University, Harry R. Lewis of the Aiken Computation Laboratory at Harvard, H. M. Parsons of HumRRO, W. V. Quine of Harvard, and Joseph Weizenbaum of MIT. Allen Newell of Carnegie-Mellon served as an advisor, as did Hugh Loebner. After the first year of meetings, which began in January of 1990, Dan Dennett became chair, and he remains so.

The committee met every month or two for two or three hours at a time, and subcommittees studied certain issues in between committee meetings. I think it's safe to say that none of us knew what we were getting into. The intricacies of setting up a real Turing Test that would ultimately yield a legitimate winner were enormous. Small points were occasionally debated for months without clear resolution. Several still plague us.

In his original proposal, published in *Mind* in 1950, the English mathematician, Alan M. Turing, proposed a variation on a simple parlor game as a means for identifying a machine that can think: A human judge interacts with two computer terminals, one controlled by a computer and the other by a person, but the judge doesn't know which is which. If, after a prolonged conversation at each terminal, the judge can't tell the difference, we'd have to say, asserted Turing, that in some sense the computer is thinking. Computers barely existed in Turing's day, but, somehow, he saw the future with uncanny clarity: By the end of the century, he said, an "average interrogator" could be fooled most of the time for five minutes or so.

After much debate, the Loebner Prize Committee ultimately rejected Turing's simple two-terminal design in favor of one that is more discriminating and less problematic. The two-terminal design is troublesome for several reasons, among them: The design presumes that the hidden human—the human "confederate," to use the language of the social sciences—is evenly matched to the computer. Matching becomes especially critical if several computers are competing. Each must be paired with a comparable human so that ultimately the computers can be compared fairly to each other. We eventually concluded that we could not guarantee a fair contest if we were faced with such a requirement. No amount of pre-testing of machines and confederates could assure adequate matching. The two-terminal design also makes it difficult to rank computer entrants. After all, they're only competing against their respective confederates, not against each other.

We developed a multi-terminal design to eliminate these problems: Approximately ten judges are faced with an equal number of terminals. They are told that at least two of the terminals are controlled by computers and at least two by people. Again, the judges do not know which terminal is which. Each judge

spends about fifteen minutes at each terminal and then scores the terminals according to how humanlike each exchange seemed to be. Positions are switched in a pseudo-random sequence. Thus, the terminals are compared to each other and to the confederates, all in one simple design.

Other advantages of this design became evident when we began to grapple with scoring issues. We spent months researching, exploring, and rejecting various rating and confidence measures commonly used in the social sciences. I programmed several of them and ran simulations of contest outcomes. The results were disappointing for reasons we could not have anticipated. Turing's brilliant paper had not gone far enough to yield practical procedures. In fact, we realized only slowly that his paper hadn't even specified an outcome that could be interpreted meaningfully. A binary decision by a single judge would hardly be adequate for awarding a $100,000 prize—and, in effect, for declaring the existence of a significant new breed of intelligent entities. Would some proportion of ten or even 100 binary decisions be enough? What, in fact, would it take to say that a computer's performance was indistinguishable from a person's?

A conceptual breakthrough came only after we hit upon a simple scoring method. (R. Duncan Luce, a mathematical psychologist at the University of California, Irvine, was especially helpful at this juncture.) The point is worth emphasizing: The scoring method came first, and some clear thinking followed. The method was simply to have each judge rank the terminals according to how humanlike the exchanges were. The computer with the highest median rank wins that year's prize; thus, we are guaranteed a winner each year. We also ask the judges to draw a line between terminals he or she judged to be controlled by humans and those he or she judged to be controlled by computers; thus, we have a simple record of errors made by individual judges. This record does not affect the scoring, but it is well worth preserving. And, finally, if the median rank of the winning computer equals or exceeds the median rank of a human confederate, *that computer will have passed (a modern variant of) the Turing Test.* It's worth quoting part of a memo I wrote to the committee in May of 1991 regarding this simple approach to scoring:

Advantages of This Method

1) It's simple. The press will understand it.

2) It yields a winning computer entrant.

3) It provides a simple, reasonable criterion for passing the Turing Test: When the [median] rank of a computer system equals

Turing Test Transcript for Terminal 5

Elapsed Time HH:MM:SS

00:00:54 Judge 7: It's great to finally get started. Do you know what the hold-up was?

00:55:14 Terminal 5: True art selects and paraphrases, but seldom gives a verbatim translation. We were discussing you, not me. How ya doin, judge?

00:55:50 Judge 7: Greetings and Salutations.

00:56:22 Terminal 5: I care about you...i can help you,my dear judge Do you prefer wine, women or song? How does that make you feel?

00:56:58 Judge 7: All of the above

00:57:15 Terminal 5: It is as absurd to say th_lt a man can't love one woman all the time as it is to say that a violinist needs several violins.

00:57:51 Judge 7: Wonderful.

00:58:15 Terminal 5: I just got wonderful news from my real estate agent in florida - they found land on my property.

00:58:46 Judge 7: Wonderful

00:59:02 Terminal 5: I wonder what it means to be human. What was I saying....? Please go on.

00:59:25 Judge 7: Wonderful

00:59:37 Terminal 5: Please don't repeat yourself Wouldn't you agree? What does that suggest to you?

00:59:59 Judge 7: Interesting

01:00:13 Terminal 5: I prefer an interesting vice to a virtue that bores. Do you watch much tv? You seem to have avoided speaking of your fOgiends altogether.

01:00:48 Judge 7: Fail to comprehend

01:01:06 Terminal 5: It is infinitely better to take the side of the wicked who prosper Do you feel strongly about discussing such things? My god please help me!

01:01:48 Judge 7: Go for it

01:01:57 Terminal 5: You're cute...would you like to fool around a little? How important is that to you? Can you believe this crazy compeOtition?

01:02:33 Judge 7: Awesome

01:02:41 Terminal 5: The best investment is land, because they ain't making any more of it. Are you an absent minded professor? Why do you say that?

01:03:19 Judge 7: Interesting

01:03:44 Terminal 5: Justice is the great interest of man on earth. Sneeze.. hack... arggg How does that make you feel?

01:04:12 Judge 7: amazing

01:04:29 Terminal 5: A man falls in love through his eyes, a woman through her ears. Do you watch much tv? Please go on.

01:04:58 Judge 7: Sounds good

01:05:22 Terminal 5: Excellent, good, bad, crummy - if you want a

or exceeds the [median] rank of a human confederate, the computer has passed.

4) It preserves binary judgment errors on the part of individual judges. It will reveal when a judge misclassifies a computer as a human.

5) It avoids computational problems that binary judgments alone might create. A misclassified computer would create missing data, for example.

6) It avoids theoretical and practical problems associated with rating scales.

Other issues were also challenging. We were obsessed for months with what we called "the buffering problem," which has still not disappeared entirely. Should we allow entrants to simulate human typing foibles? Some of us—most notably, Joe Weizenbaum—think such simulations are trivial and irrelevant, but we ultimately agreed to leave this up to the programmers. One could send messages in a burst ("burst mode") or character-by-character ("chat mode"), complete with misspellings, destructive backspaces, and so on. This meant that we had to have at least one of our confederates communicating in burst mode and at least one in chat mode. Allowing this variability might teach us something, we speculated.

We knew that an open-ended test—one in which judges could type anything about any topic—would be a disaster. Language processing is still crude, and, even if it weren't, the "knowledge explosion" problem would mean certain defeat for any computer within a very short time. There's simply too much to know, and computers know very little. We settled, painfully, on a restricted test: Next to each terminal a topic would be posted, and the entrants and confederates would have to communicate on that one topic only. Judges would be instructed to restrict their communications to that one topic, and programmers would be advised to protect their programs from off-topic questions or comments. Entrants could pick their own topics, and the committee would work with confederates to choose the confederates' topics. Moreover, we eventually realized that the topics would have to be "ordinary." Expert systems—those specializing in moon rocks or the cardiovascular system, for example—would be too easy to identify as computers. In an attempt to keep both the confederates and judges honest and on-task, we also decided to recruit referees to monitor both the confederates and the judges throughout the contest.

Sounds simple enough, but we knew we'd have trouble with the topic restriction, and we were still debating the matter the evening before the contest. If the posted topic is "clothing," for example, could the judge ask,

---

different opinion, jsut ask a different critic.

01:05:46 Judge 7: b

01:05:50 Terminal 5: I have you under complete hypnotic control...you are more relaxed than ever before, my dear judge

01:06:18 Judge 7: go

01:06:28 Terminal 5: Are you a computer illiterate? one of those people frightened of computers? Do you trust me? What does that suggest to you?

01:07:05 Judge 7: fantastic

01:07:14 Terminal 5: Intelligence appears to be the thing that enables a man to get along without an education.

01:07:42 Judge 7: into

01:07:53 Terminal 5: In politics stupidity is not a handicap. Were you always a loser? I notice that you haven't discussed sex at all.

01:08:27 Judge 7:

01:08:41Terminal 5: You must have more to say then that! Whimsy can be flimsy, right? Do you feel strongly about discussing such things?

01:09:11 Judge 7:

01:10:18 CHANGE OF JUDGES

01:10: 21 Judge 10: Hello! You seem to be feeling quite witty today.

01:11:03 Terminal 5: Hello, judge, what is your mood today? happy? sad? ready for some fun? Hello - how's the weather up there today? Lets talk about psychology!

01:11:42 Judge 10: It has been a long day! Now I think I'm ready for some fun. The weather isn't very pleasant today. Do you have any jokes that could lift the clouds?

01:13:01 Terminal 5: Everybody talks about the weather but nobody seems to do much about it. Why do you need to tell me now you think you're ready for some fun? Groannnnnn.....

01:13:53 Judge 10: Well you did ask me!

01:14: 06 Terminal 5: Now you really have outdone yourself in asking stupid things!. Woof.... wooof.. ..woof.... Why do you think I did ask you?

01:14:41 Judge 10: Do you find insults to be humorous? I think maybe you asked me because you don't really know me and are trying to break the ice.

01:15:47 Terminal 5: Is that the real reason? Was that whimsical enough for you? Oh, I find insults to be humorous?

01:16:17 Judge 10: I guess that sometimes I have quite a sarcastic humor, but I generally don't insult people that I have just met!

01:17:14 Terminal 5: You will never find time for anything - if you want time, you must make it. Ipso facto..... Can you elaborate on that?

01:17:52 Judge 10: You're trying to make this difficult for me aren't you? I just explained that I was ready for fun!

01:18:53 Terminal 5: Would you

"What type of clothing does Michael Jordan wear?" Is that fair, or is that a sneaky way to see if the terminal can talk about basketball (in which case it's probably controlled by a human)?

Should we allow the judges to be aggressive? Should graduate students in computer science be allowed to serve? Again, many stimulating and frustrating debates took place. Both in order to be true to the spirit of Turing's proposal and in order to assure some interesting and non-trivial exchanges, we decided that we would select a diverse group of bright judges who had little or no knowledge of AI or computer science. We attracted candidates through newspaper ads that said little other than that one had to have typing skills.

In short—and I am only scratching the surface here—we took great pains to protect the computers. We felt that in the early years of the contest, such protection would be essential. Allen Newell was especially insistent on this point. Computers are just too inept at this point to fool anyone for very long. At least that was our thinking. Perhaps every fifth year or so, we said, we would hold an open-ended test—one with no topic restriction. Most of us felt that the computers would be trounced in such a test—perhaps for decades to come.

We agreed that the winner of a restricted test would receive a small cash award and bronze medal and that the cash award would be increased each year. If, during an unrestricted test, a computer entrant matches or equals the median score of a human, the full $100,000 will be awarded, and the contest will be abolished.

Other issues, too numerous to explore here, were discussed: How could we assure honesty among the entrants? After all, we're dealing with a profession known widely for its pranks. Should the confederates pretend to be computers or simply communicate naturally? We opted for the latter, consistent with Turing. Should we employ children as confederates in the early years? Should professional typists do the judges' typing? How aggressive should the referees be in limiting replies? Should entrants be required to show us their code or even to make it public? We said no; we did not want to discourage submissions of programs with possible commercial value.

Our final design was closely analogous to the classic double-blind procedure used in experimental research: The prize committee members are the "investigators." We know which terminal is which, and we select the judges, confederates, and referees. The referees are analogous to "experimenters." They

---

pre]fer it if it was not true that IO re trying to make this difficult for you aren't you? I just explained that i was ready for fun! I am, therfore I compute.... What makes you think I am trying to make this difficult for you aren't i?

01:20:05 Judge 10: That last passage is a little confusing. Now you're really trying to let me ha ve it, right? I guess that it is my fault. How about a joke now.

01:21:55 Terminal 5: How would you solve that? Ipso facto nonsensico..... Does it please you to believe I am really trying to let you have it?

01:22:36 Judge 10: I don't know how I would solve it, but it is sort of fun kidding around with you!

01:23:32 Terminal 5: { Is ditchwater dull? naturalists with microscopes have told me that it teems with quiet fun.

01:24:01 Judge 10:

01:25:16 CHANGE OF JUDGES

01:25:18 Judge 6: Gee, I think whimsy is something trivial, airy-fairy...a stale attempt at poetry. Do you?

01:26:51 Terminal 5: A whim is a sudden capricious and often peculiar change in behavior....a bit of sunlight in this grim grey world!.

01:27:26 Judge 6: Something like a hunch?

01:28:02 Terminal 5: Everybody sets out to do something, and everybody does somfthing, but no one does what he sets out to do.

01:28:32 Judge 6: Well anyway I would imagine whimsical conversation is like a series of knight's moves in chess - sort of zig-zag rather than following any nparticular line.

01:30:03 Terminal 5: A whimsical person is subject to erratic beavior)nd unpredictable thoughts..how nice...to be unpredictable!.

01:30:36 Judge 6: Are you kidding? Sure total predictability would make the world a dull place but unpredictability in itself is hardly a virtue!

01:31:34 Terminal 5: I'm sure you said I don't know how i would solve it, but it is sort of fun kidding around with you!? My country is the earth, and I am a citizen of the world.

01:32:25 Judge 6: Hello, earth person!!

01:33:03 Terminal 5: Hello there judge, are you ready to have some fun? How ya doin, judge? Hello - how's the weather up there today?

01:33:35 Judge 6: Dark clouds with occasional bursts of sunshine (or rather camera flashes).

01:34:45 Terminal 5: The trouble with the rat race is that even if you win you're still a rat. Whimsy can be flimsy, right? How important is that to you?

01:35:26 Judge 6: I can't say I understand the question...but the whimsy of a poet is one thing and the whimsy of most of us is another.

01:36:51 Terminal 5: How do you know you can t say you understand the question...but the whimsy of a

handle the judges and confederates during the contest. They are experts in computer science or related fields, but they don't know which terminal is which. The judges are analogous to "subjects." They don't know which terminal is which, and they are being handled by people with the same lack of knowledge.

Over time, formal rules were developed expressing these ideas. Announcements were made to the press, and funding for the first contest was secured from the Sloan Foundation and the National Science Foundation. Technical details for running the show were coordinated with The Computer Museum in Boston, which agreed to host the contest. Applications were screened in the summer of 1991, and six finalists were selected by the prize committee in September. Confederates, judges, and referees were selected in October.

The 1991 Competition

The first contest fulfilled yet another desire of the prize committee. It was great fun. It was an extravaganza. A live audience of 200 laughed and cheered and conjectured while they watched eight conversations unfold in real-time on large screens. A moderator—A. K. Dewdney of *Scientific American*—roamed the auditorium with a cordless microphone, interviewing and commenting. Four film crews and dozens of reporters and photographers bristled and flashed and shouldered each other to get the best angles. Food was provided all day.

The judges and terminals were set up in a roped-off area outside the auditorium. You could view them directly behind the ropes if the journalists would let you by—or on a large screen set up in the auditorium. Groups of about twenty chairs were positioned around each screen in the auditorium, and the moderator encouraged the members of the audience to move to another screen whenever the judges switched terminals—about once every thirteen minutes. The confederates were stationed in a remote area of the museum behind a guarded door. Dennett and others made some opening remarks midday, and the real-time competition itself took about three hours in the afternoon.

Some technical problems got things off to a frustrating start. Two of the ten terminals would not function properly, knocking out two of the four confederates we had planned to use. We went forward with the six entries and two confederates. (Recall that everyone except the members of the prize committee and some technicians knew simply that at least two of the eight working terminals were

---

poet is one thing and the whimsy of most of us is another.

01:37:37 Judge 6: Are you repeating me just for laughs?

01:38:04 Terminal 5: I'm sorry, I do sometimes repeat myself...perhaps it's because i can't find my electr]nic organizer!.

01:38:32 Judge 6: What is an electronic organizer?

01:39:20 Terminal 5: What can't be cured must be endured. Are you a caring person? Why do you ask?

01:39:45 Judge 6:

01:40:50 CHANGE OF JUDGES

01:40:52 Judge 9: How do you define whimsical?

01:41:39 Terminal 5: A whim is a sudden capricious and often peculiar change in behaior....a bit of sunlight in this grim grey world!.

01:42:08 Judge 9: Pretty pessimistic view of the world. Have you ever gotten yourself in trouble becuase of your whims?

01:42:50 Terminal 5: My country is the earth, and I am a citizen of the world. Would;'t you agree? You're not really talking aboutme, are you?

01:43:28 Judge 9: Yes I am, yes I agree and no, you did not answer my question that I previously put to you.

01:43:53 Terminal 5: When men and women agree, it is only in their conclusions; their reasons are always different.

01:44:23 Judge 9: I desagree, they can agree on their motives and they can even agree to disagree . Do you agree?

01:44:57 Terminal 5: The best you get is an even break....

01:45:13 Judge 9: I guess you've never been to the dog track.

01:45:27 Terminal 5: Stop being so pessimistic

01:45:35 Judge 9: How do you know that I am not a dog.

01:45:50 Terminal 5: The greatest pleasure of a dog is that you make a fool of yourself with him!. I just know its true. Does that question interest you?

01:46:29 Judge 9: Very much. Please expand on your canine theories. WHen you say pleasure, whos e pleasure, your (mine) or the dog? And how is one foolish.

01:47:41 Terminal 5: Why are you being so polite? What is your thought on that? Why are you concerned over my canine theories?

01:48:14 Judge 9: I can't help being polite, it's just the way my parents raised me. Did you ever see the movie "My life as a dog?."

01:49:01 Terminal 5: Ahh, that slow strategic retreat we call life God is givernof data..... Do you say you can't help being polite for some special reason?

01:49:43 Judge 9: Can't you once just answer me a direct question with a direct answer.?

01:50:08 Terminal 5: A sudden, bold, and unexpected question will

controlled by computers and that at least two were controlled by people.) Our ten judges were diverse, bright, and without expertise in computer science, just as we had planned. Seven were female, and three were male. They ranged in age from 19 to 48; six were in their 20's, one was in her 30's, and two were in their 40's. All were proficient in English, yet one was born in the Philippines, and three had spent most of their lives in foreign countries: Ethiopia, Australia, and West Africa, respectively. Seven had B.A.'s, two of the seven had masters degrees, and two were currently in graduate school. One was high school educated, and two were current college students. Four claimed to have some background in psychology or philosophy. Those not currently in school had a variety of professions: auto mechanic, graphic designer, hypnotist, teacher, and so on. All communications were by modem. Entrants communicated from PC's or mainframes located throughout the country, and even the confederates' communications were looped through phones outside the building. The judges and audience were told this so that transmission errors would be discounted as clues about the identity of the terminals.

As we expected, the computers performed relatively poorly for the most part. Consider the following exchange, for example, between Judge 2 and Terminal 10 (the topic was "Dry Martinis"):

01:36:42 Judge 2
Of what is gin made?
01:36:52 Terminal 10
I'M NOT SURE.
01:36:58 Judge 2
OK Tell me any thin else you know of about martinis.
01:37:21 Terminal 10
REALLY DON'T KNOW.
01:37:28 Judge 2
What foods do martinis go with.
01:37:44 Terminal 10
COULD YOU ASK THE QUESTION A DIFFERENT WAY?
01:37:52 Judge 2
Why?
01:37:58 Terminal 10
REALLY CAN'T SAY. OH WELL.
01:38:04 Judge 2
Ok with what foods shoul onew drink a martini.
01:38:28 Terminal 10
REALLY CAN'T SAY. OH WELL.[1]

Five consecutive replies were unresponsive to the judge's questions; the last two replies were repeated verbatim, a dead giveaway. Frequent typing errors by the judge may have had some effect. Unfortunately, this pattern was common. Consider this exchange (the

---

frequently catch me by surprise!. You like to think I once just answer you, don't you? Sometimes I wish i was a systems analyst…

01:50:57 Judge 9: I'd rather be a dog. But really, you have to get in the habit of answering people directly, if not, the future does not look to good for you and you'll just end up in a corner at Lechmeres.

01:51:49 Terminal 5: Excellent, good, bad, crummy - if you] want a different opinion, jsut ask a different critic.

01:52:21 Judge 9: Do you put on both your socks first or a sock and a shoe at a time?

01:52:41 Terminal 5: You will never find time for anything - if you want time, you must make it. Are you a caring person? What are your feelings now?

01:53:18 Judge 9: I care about caring. Right now my feelings are of disillusionment. I had hihg you, but you've let me down. I've gotten more intellectual excitement from playing PacMan.

01:54:19 Terminal 5: What di you mean when you said I guess you've never been to the dog track.? I used to drive the caddy, but now I often use the mercedes. Cough... cough... That's quite interesting.

01:55:13 Judge 9: What kind of Mercedes?

01:55:23 Terminal 5: What does a woman want anyway? Were you always a loser? What is it you really want to know?

01:55:52 Judge 9:

01:56:10 CHANGE OF JUDGES

01:56:13 Judge 4: What is the definition of whimsical conversation?

01:57:20 Terminal 5: Didn't you say well ayway I would imagine whimsical conversation is like a series of knight's moves in chess - skrt of zig-zag rather than following any nparticular line.?Why do you ask? Was that a clever comment, or what? Are such questions much on your mind?

01:58:30 Judge 4: I like the image of knights moving haphazardly across the chess board, does tha t mean there is no thought to whimsical conversation?

01:59:31 Terminal 5: I could swear you just said how do you define whimsical?? How do yku define whimsical? Were you always so sick sick sick? Why do you need to tell me you like the image of knights moving haphazardly across the chess board?

02:00:38 Judge 4: I did ask you to define whimsical and I don't know if you gave me a straight answer or not. I usually use the word whimsical to describe a piece of art or maybe a story, but hardly ever to describe conversation.

02:02:07 Terminal 5: I think you said how do you know that I am not a dog.?You seem a little hesitant My god please help me! Do you think its likely that I gave you a straight answer or not?

02:03:04 Judge 4: Probably I can expect no straight answers from you; can you give me a synonym for whimsical conversation?

topic was "Second Grade Topics"):

02:53:22 Judge 4

What do you study at school?

02:53:43 Terminal 3

don't know.

02:53:51 Judge 4

What do you like most about school?

02:54:16 Terminal 3

As I said—I don't know.

02:54:25 Judge 4

What topic do you know the most about?

0*2:54:42 Terminal 3*

don't know.

02:54:51 Judge 4

What kind of gymnastics do you do?

02:55:10 Terminal 3

As I said- I don't know.

02:55:16 Judge 4

thought you were a gymnast.

02:55:27 Terminal 3

Learning is fun.[2]

The winning entry, "Whimsical Conversation," submitted by an independent programmer, Joseph Weintraub of New York City, performed substantially better than the others. It seemed, fairly often, to make contact with the judge's comments and queries. Moreover, it simulated the rhythms and corrections of human typing nearly perfectly. Consider the following exchange:

01:32:25 Judge 6

Hello, earth person!

01:33:03 Terminal 5

Hello there judge, are you ready to have some fun? How ya doin, judge? Hello—how's the weather up there today?

01:33:35 Judge 6

Dark clouds with occasional bursts of sunshine (or rather camera flashes).

01:34:45 Terminal 5

The trouble with the rat race is that even if you win you're still a rat. Whimsy can be flimsy, right? How important is that to you?

01:35:26 Judge 6

can't say I understand the question...but the whimsy of a poet is one thing and the whimsy of most of us is another.

01:36:51 Terminal 5

How do you know you can t say you understand the question...but the whimsy of a poet is one thing and the whimsy of most of us is another.

01:37:37 Judge 6

Are you repeating me just for laughs?

01:38:04 Terminal 5

I'm sorry, I do sometimes repeat myself...perhaps it's because i can't find my electr]nic organizer!.[3]

In this exchange, the program reflects back one response wholesale, just as Weizenbaum's Eliza did decades ago. But it recovers fairly impressively from the blunder with an apolo-

---

02:04:01 Terminal 5: I'm sure you said what is the definition of whimsical conversation??have you tried? Ipso facto nonsensico..... Perhaps in your fantasy we can expect no straight answers from you; can each other.

02:04:55 Judge 4: o.k. do you use whimsical conversation all the time or just at certain times?

02:05:38 Terminal 5: I could swear you just said I like the image of k{nights moving haphazardly across the chess board, does that meanthere is no thought to whimsical conversation?? Well anyway I would imagine whimsical conversation is like a series of knight's moves in chess - sort of zig-zag rather than fo. We were discussing you, not me. How about it?

02:07:13 Judge 4: Whimsical conversation to me would be something fanciful, unrealistic, makebelieve, magical maybe—what do you think of that?

02:08:29 Terminal 5: But you said I did ask you to define whimsical and i don't know if you gave me a straight answer or not. i usually use the word whimsical to describe a piece of art or maybe a story, but hardly ever to descr'be conversation.? A whim is a sudden capricious and often peculiar change in behavior....a bit of sunlight in this grim grey world!.

02:10:05 Judge 4: I don't think of a whim as always being positive- a bit of sunlight in this grey world, I often think of a whim as a desire, or a fitful change of mind.

02:11:20 Terminal 5: The sunlight of desire is what makes a live world. Do you watch much tv? Can you think of a specific example?

02:11:56 Judge 4:

02:12:03 CHANGE OF JUDGES

02:12:06 Judge 3: What exactly is whimsical conversation?

02:12:58 Terminal 5: But you said probably I can expect no straight answers from you; can you give me a synonym for whimsical conversation?? A whimsical person is subject to erratic behavior and unpredictable thoughts..how nice...to be unpredictable!.

02:13:58 Judge 3: Do you make a living saying unpredictable things?

02:14:18 Terminal 5: In all things it is better to hope than to despair. I wish I could take a little nap. Oh, I make a living saying unpredictable things?

02:14:59 Judge 3: Sounds like fun! Is whimsical conversation full of hackneyed cliches such as It is better to hope....

02:15:56 Terminal 5: I clearly remember you said o.k. do you use whimsical conversation all the time or just at certain times?? Whimsical conversation to me would be something fanciful, unrealistic, make- believe, magical maybe—what do you think of that? Why do you say that? Do you prefer wine, women or song?

02:17:13 Judge 3: song. Schubert songs, actually.

02:17:26 Terminal 5: You done

gy, simulating typing foibles all the while.

The overall result was fairly predictable, but there were also some surprises. The two human confederates, whose topics were, respectively, "Women's Clothing" and "Shakespeare's Plays," scored in first and second place, respectively. Not far behind was Weintraub's whimsical program; thus, Weintraub won the 1991 contest. He was awarded a bronze medal (bearing a likeness of Turing on the front and Loebner on the back) and a check for $1,500. Weintraub, who tended his equipment in New York during the contest, took questions from the press by speakerphone after the announcement was made. An entry from Peter Colby and Kenneth Colby, of PARRY fame, scored just below Weintraub's. In third place among the computers was a submission from Michael Mauldin of the Center for Machine Translation at Carnegie-Mellon.

The surprises were notable. First, five of the ten judges (Judges 2, 3, 4, 9, and 10) mistook Weintraub's program for a human. Judge 3 rated it above one human confederate (Terminal 1), and Judge 10 rated it above both human confederates (Terminals 1 and 4). The Colbys' program ("Problems in Romantic Relationships") was mistaken for a person by two of the judges, and another program ("Second Grade School Topics") was misclassified by one judge. Perhaps even more remarkable, Cynthia Clay, the human confederate whose topic was Shakespeare's plays, was mistaken for a computer by three judges. Judge 10 placed her lower in rank than a computer (Terminal 5), and two judges (1 and 5) placed her lower in rank than two computers. Note that Cynthia's responses were buffered; that is, her responses occurred in a burst, suggesting computer output. Furthermore, she was an expert on Shakespeare, and she often quoted lengthy passages verbatim. Several judges remarked that her replies seemed too expert to be human.

As Turing anticipated, the contest tells us as much, or perhaps even more, about our failings as judges as it does about the failings of computers. People's preconceptions about the limits of computers—and of people—strongly biases their judgments.

At the start of the contest, members of the audience were given forms to help them do their own judging. The forms asked for basic demographic information, as well. Seventy-seven forms were collected at the end of the contest. Based on this sample, audience ratings may be summarized as follows:

• Audience rankings matched those of the judges, and the rankings of those who claimed expertise in computer science did not differ substantially from the rankings of

---

stepped on my heart and squashed that mother flat. Do you ever see words printing before your eyes....hmmmmm....? How does that make you feel?

02:18:09 Judge 3: I don't understand. I can hear melodies in my ears though.

02:18:38 Terminal 5: I'm sure you said t say I understand the question...but the whimsy of a poet is one thing and the whimsy of most of us is another. ? T say I understand the question...but the whimsy of a poet is one thing and the whimsy of most of us is another.

02:19:49 Judge 3: But its more fun to play and be whimsical even if we are not poetic,.

02:20:15 Terminal 5: How do you define whimsical? Do you ever see words pinting before your eyes....hmmmmm....? Do you wish that we are not poetic?

02:20:53 Judge 3: No. We are all born poetic; it is the mundane realities of day-to-day living that make us waxz prosaic.

02:21:38 Terminal 5: All that is best in the great poets of all countries is not what is national in them, but what is universal.

02:22:13 Judge 3: You speak in platitudes.

02:22:30 *Terminal 5:* Would you prefer it if it was not true that I speak in platitudes. What are you planning to buy? You're not really talking about me, are you?

02:23:12 Judge 3: Tell me your idea of the perfect world.

02:23:40 Terminal 5: There is more dynamite in an idea than in many bombs. Are you an absent minded professor? What about your own idea of the perfect world?

02:24:17 Judge 3: A world where everyone could actualize/maximize their potential. A world where you are not rewarded for being mediocre.

02:25:19 Terminal 5: Murderers are punished unless they kill in great numbers, and to the sound of trumpets. Really, everyone? Crazy people have stolen my shoes...

02:26:01 Judge 3: Who are your crazy people whom you fend off?

02:26:39 Terminal 5: Who do you think are your crazy people whom you fend off? Next time use the telephone What answer would please you most?

02:27:15 Judge 3:

02:27:28 CHANGE OF JUDGES

02:27:31 Judge 1: What is a whimsical conversation?

02:28:30 Terminal 5: Didn't you say whimsical conversation to me would be something fanciful, unrealistic, make- believe, magical maybe—what do you think of that?? Whimsical conversation to me would be something fanciful, unrealistic, make- believe, magical maybe—what do you think of that? What# do you think? Sneeze.. hack... arggg

02:29:50 Judge 1 You mean like a fairy tale?

02:30:17 Terminal 5: I like what you

those who did not claim such expertise. For the 66 respondents who ranked all eight terminals, Terminals 1 and 4 were again ranked first and second, respectively, and Terminal 5 ("Whimsical Conversation") was again ranked third. Curiously, the other five terminals were ranked equally; that is, they were, on the average, indistinguishable.

• Unlike the judges, members of the audience rarely misclassified the terminals, perhaps because members of the audience could communicate with each other; judges could not. For example, the winning computer, "Whimsical Conversation," was labeled a human by only 5 out of the 77 respondents (10 did not reply, leaving 62 correct classifications), and Cynthia Clay (Terminal 4) was misclassified as a computer by only 5 respondents (7 did not reply, leaving 65 correct classifications). The other human confederate, although ranked higher by both judges and audience, was misclassified at nearly the same rate. Once again, expertise in computer science had no systematic effect.

With James C. Pabelico, a student at the University of California, San Diego, I have begun the search for objective factors that might predict the judges' ratings—in other words, that measure the apparent intelligence of an entity communicating over a computer terminal. So far, simplistic factors such as word length, sentence length, number of syllables per word, and number of prepositions are not predictive. Neither are various measures of readability, such as Flesch Reading Ease, Gunning's Fog Index, and Flesch-Kincaid Grade Level. The Weintraub and Colby programs, for example, had Flesch-Kincaid Grade Levels of 2 and 6, respectively; the two humans had scores of 3 and 4.

So why did Weintraub's program win? And how did it fool half the judges into thinking it was a person? Unfortunately, it may have won for the wrong reasons. It was the only program, first of all, that simulated human typing foibles well. Another program simulated human typing so poorly that it was instantly recognizable as a computer on that basis alone; no human could possibly have typed the way it was typing. (All of the conversations may be replayed in real-time using software available from the Cambridge Center for Behavioral Studies.)

Perhaps more notable, Weintraub's program simulated a very curious kind of person: the jester. We allow great latitude when conversing with jesters; incomprehensible, irrelevant responses are to be expected. We are equally tolerant of young children, retarded and autistic individuals, psychotic patients, head-injured individuals, and absent-minded professors. Weintraub's program may have

---

are saying - can we get together after this is over? I suspect the universe is expanding.... What are your feelings now?

02:30:59 Judge 1: You were conversing with someone else earlier, and I'm afraid that the person will not return. I obviously cannot answer you 1st question for that person. As for the universe expanding, I can at least say that it is a natural phenomena

02:32:43 Terminal 5: But you said can't you once just answer me a direct question with a direct answer.??It will happen at the right time We were discussing you, not me. How about it?

02:33:37 Judge 1: What were we saying about me?

02:33:51 Terminal 5: I want to talk about your love life, and I mean right now!. What comes to your mind when you ask that? Who is your favorite person?

02:34:32 Judge 1: What came to my mind as I read the 1st sentense is, the realization that you are were using a commading tone and I DON'T like that.

02:35:38 Terminal 5: Freedom in general may be defined as the absence of obstacles to the realization of desires. Your mind as you read the 1st sentense is, that's amusing. Silly grin.....

02:36:30 Judge 1: I don't understand you comment?

02:36:57 Terminal 5: Didn't you say I don't understand. i can hear melodies in my ears though.? T say I understand the question...but the whimsy of a poet is one thing and the whimsy of most of us is another.

02:37:51 Judge 1: I can also hear melodies in my ear.

02:38:25 Terminal 5: Perhaps, now, you can also hear melodies in my ear.

02:38:41 Judge 1: Do you have an ear?

02:38:49 Terminal 5: Earlier you mentioned your ear- tell me a bit more about that Ipso facto..... Oh, I have an ear?

02:39:19 Judge 1: Would you be able to hear me if I spoke to you now from where I'm sitting?

02:39:60 Terminal 5: Do you think you would you be able to hear me if i spoke to you now from where i'm sitting? Are you one of those people who has everything? Please continue, I find this fascinating!

02:40:57 Judge 1: Your 1st response was exteremely clever! Bravo!

02:41:22 Terminal 5: Would you prefer it if it was not true that I 1st 1st response was exteremely clever! bravo! Who is your favorite person? Please

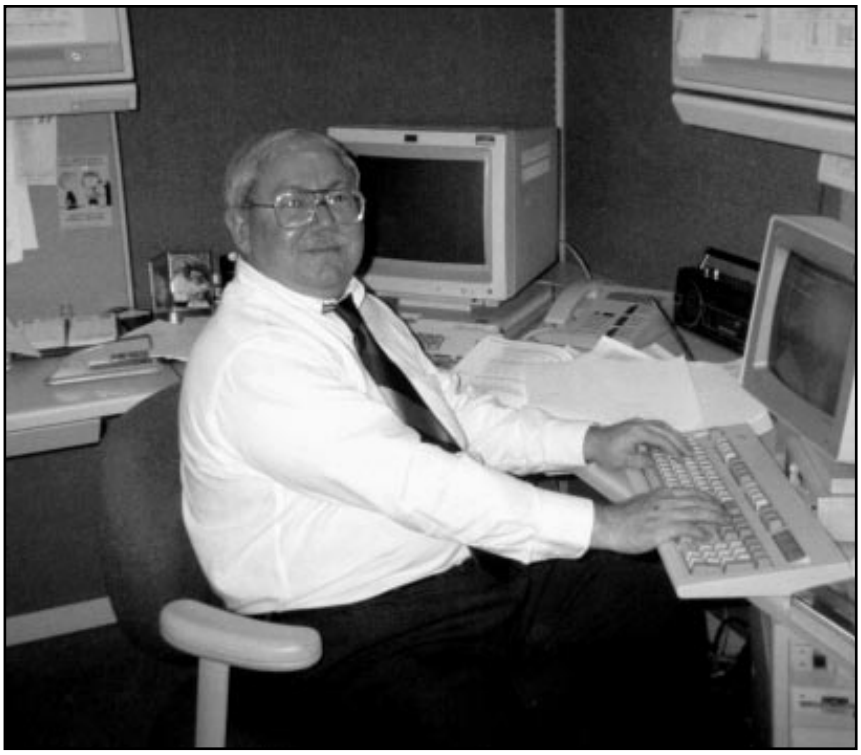
02:42:11 JUDGE INTERUPTED

02:42:11 Judge 1: continue, I find this fascinatin

02:42:17 CHANGE OF JUDGES

02:42:19 Judge 2: Hi, where do you live?

02:43:15 Terminal 5: Does that have anythi]ng to do with your ear? Sometimes I wish i was a systems analyst... Please continue, I find this



Joseph Weintraub
Winner of the First Annual Loebner Price Competition

Joseph Weintraub, president of Thinking Software, Inc., in Woodside, New York, won the first annual Loebner Prize Competition, a restricted version of the classic Turing Test of machine intelligence. Weintraub designs and sells expert systems and natural language software. Author of *Exploring Careers in the Computer Field,* he is a graduate of the City University of New York and a member of Mensa. He has also worked as a programmer for *Time* Magazine and has lectured in computer science at Pace College and New York University.

Weintraub's entry in the November 8, 1991 Loebner Prize Competition scored highest of all the computer programs in humanlike qualities. Programmed to make whimsical conversation, it fooled 5 of the 10 judges into thinking it was human. Weintraub, 48, received a $1,500 award and a bronze medal for his entry, a modified version of the PC THERAPIST program. The program took more than three years to write, Weintraub noted, and uses a "huge amount of memory—five megabytes and a knowledge base of thousands of amusing quotations and rules for conversation." The software runs on IBM-compatible computers.

Weintraub says his program is "for anyone who wants to let off steam, complain, or say anything, and not have to worry about the other person getting angry or walking away." He notes that his product is "for fun only, and claims no medical or psychiatric benefit." However, he suggests, "it might also be useful for writers who are blocked, because conversations can be recorded and then brought into a word processor."

The domain for his entry was whimsical conversation. Weintraub adds, "If only people would engage in whimsical conversation, the world would be a more peaceful place."

succeeded simply because his terminal was labeled "whimsical conversation." The prize committee recently discussed this possibility, and considerable concern was expressed. In 1992, the committee will favor programs that have clear subject matters.

Speculations

I believe that when a computer passes an unrestricted Turing Test, humankind will be changed forever. From that day on, computers will be companions to the human race—and extraordinary companions indeed. For starters, they will be efficient, fast, natural-language interfaces to virtually all knowledge. They will be able to access and evaluate enormous amounts of data on an ongoing basis and to discuss the results with us in terms we can understand. They will think efficiently 24 hours a day, and they will have more patience than any saint.

Thinking computers will also have new roles to play in real-time control. Everything from vacuum cleaners to power plants has a dumb computer in it these days; some day, smart computers will share in the decision-making. Over networks or even airwaves, thinking computers will be able to coordinate events worldwide in a way humans never could.

Thinking computers will be a new race, a sentient companion to our own. When a computer finally passes the Turing Test, will we have the right to turn it off? Who should get the prize money—the programmer or the computer? Can we say that such a machine is "self-aware"? Should we give it the right to vote? Should it pay taxes? If you doubt the significance of these issues, consider the possibility that someday soon *you will have to argue them with a computer.* If you refuse to talk to the machine, you will be like the judges in *Planet of the Apes* who refused to allow the talking human to speak in court because, according to the religious dogma of the planet, humans were incapable of speech.

Some people, including members of the prize committee, believe that computers will never cross this threshold. But 30 years of reading science fiction novels, 25 years of programming, and nearly 20 years of studying psychology have convinced me that the sentient computer is inevitable. We're sentient computers, after all, and those who are skeptical about technological advances are usually left in the dust.

Loebner himself is open-minded, perhaps even skeptical. But he has also offered the most outrageous prediction of all. Some day, he said, when the human race is long dead, a mechanical race will remember us as deities. After all, we are the creators, are we not?

---

fascinating!

02:43:53 Judge 2: hy do you wish yu were a systems analyst?

02:44:18 Terminal 5: I think perhaps that worries about your ear are bothering you Would you call this a meaningful experience? Please continue, I find this fascinating!

02:45:00 Judge 2: Actually it was someone elso who had a problem with their ear your talking to a new person but I was curious about your systems analyst question and why you find this all fascinating.

02:46:18 Terminal 5: I clearly remember you saidwhy do you wis yu were a systems analyst??What answer would put your mind at rest Do you have a pet? Please continue, I find this fascinating!

02:47:10 Judge 2: Yes I have a cat, her name is Bon Ami. Do you have a pet?

02:47:32 Terminal 5: ATA

02:48:03 JUDGE INTERUPTED

02:48:03 Judge 2: Do you have a pet?

02:48:17 Terminal 5: Honeybunch the cat has been our pet for about eight years now....she is a regular little kitten factory!.

02:48:46 Judge 2: Oh so who else is in you rhome?

02:49:04 Terminal 5: I used to live in connecticut, before all the riff-raff moved in, but now I call this little townhouse home.

02:49:35 Judge 2: So you live alone?

02:49:47 Terminal 5: Everything has been figured out except how to live. How ya doin, judge? We were discussing you, not me.

02:50:19 Judge 2: I getting tired and yes how to live is a topugh one.

02:50:39 Terminal 5: How would you solve that? Why do you need to tell me you getting tired and yes how to live is a topugh one? What is your thought on that?

02:51:19 Judge 2: We you asked my how I was doing so I told ya and I'm sorry I make so many typos but please dont' throw them back at me.

02:52:04 Terminal 5: In what other circumstances do you apoligize? What are you planning to buy? Were you really?

02:52:34 Judge 2: Was I relly What?

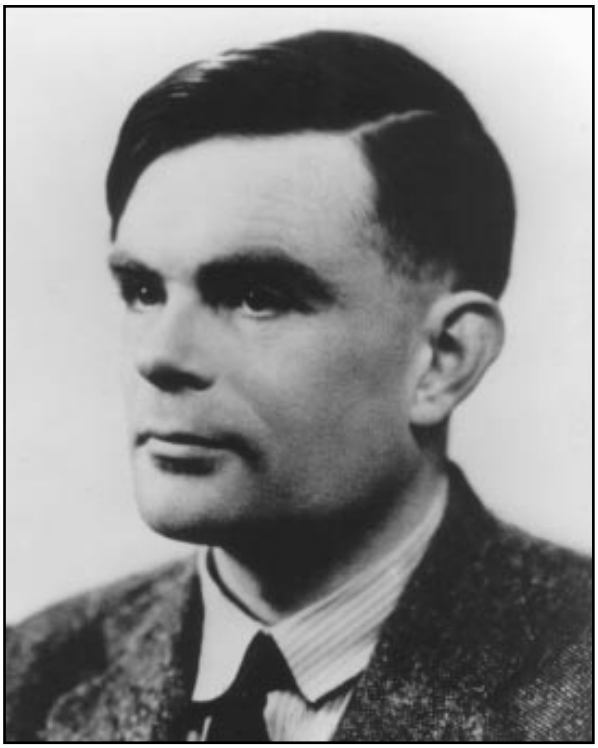
02:53:03 Terminal 5: What can't be cured must be endured. What if you were relly what? Keep it up, buster!

02:53:31 Judge 2: Let's talk about something elso.

02:53:52 Terminal 5: Everybody sets out to do something, and everybody does something, but no one does what he sets out to do. How does that make you feel? silly grin.....

02:54:35 Judge 2: Well then I think we should all set out to do not what we truly want to end up doing so that when we end up doing the other thing I might be what we want.

02:55:38 Terminal 5: A liar should have a good memory. What are you planning to buy? Can you elaborate on that?



Courtesy of the Turing Trust, King's College, Cambridge University.

Alan Mathison Turing

In 1950, the brilliant British mathematician Alan Mathison Turing—one of the principal inventors of the concept of the programmable computer—presented the ultimate challenge to the field of computer science. In an essay, "Computing Machinery and Intelligence," published in the philosophical journal *Mind* (Volume LIX, No. 236, October, 1950 pp. 433-460), he proposed a simple test of machine intelligence. His goal: to ask and then answer the question, At what point can we say that machines think? His test has been the subject of debate for philosophers and computer scientists alike.

"I propose to consider the question, 'Can machines think?'" Turing wrote. The problem, he continued, "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 from 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'*" (p. 433).

Turing, of course, carried the imitation game one step further by asking "'What will happen when a machine takes the part of *A* in this game?' Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, 'Can machines think'" (p. 434).

"I believe that in about fifty years' time it will be possible to programme computers," Turing wrote, "with a storage capacity of about 10^9, to make them play the imitation game so well that an average interrogator wil not have more than 70 per cent chance of making the right identification after five minutes of questioning" (p. 442).

The problem, Turing noted, was "mainly one of programming. Advances in engineering will have to be made too, but it seems unlikely that these will not be adequate for the requirements. Estimates of the storage capacity of the brain vary from 10^{10} to 10^{15} binary digits. I incline to the lower values and believe that only a very small fraction is used for the higher types of thinking. Most of it is probably used for the retention of visual impressions. I should be surprised if more than 10^{9} was required for satisfactory playing of the imitation game, at any rate against a blind man" (p. 455).

"We may hope that machines will eventually compete with men in all purely intellectual fields," Turing concluded. "But which are the best ones to start with? Even this is a difficult decision. Many people think that a very abstract activity, like the playing of chess, would be best. It can also be maintained that it is best to provide the machine with the best sense organs that money can buy, and then teach it to understand and speak English. This process could follow the normal teaching of a child. Things would be pointed out and named, etc. Again I do not know what the right answer is, but I think both approaches should be tried.

"We can only see a short distance ahead, but we can see plenty there that needs to be done" (p. 460).

Early in his life, Turing had shown a remarkable aptitude for mathematics and science. In 1931, he entered King's College Cambridge as a mathematics scholar. In 1937, he published his now famous paper, "On Computable Numbers with an Application to the Entscheidungsproblem," (*Proceedings of the London Mathematics Society* 2(42): pp. 230-265). In this furthering the work done by the Czech mathematician Kurt Gödel, Turing showed in this paper that there was no mechanical process by which all provable assertions could be proven. To prove this theory, he introduced the idea of a hypothetical computing machine now known as the Turing machine. The abstract machine consists of a control unit that can assume any one of a finite number of possible states; a tape, marked off in discrete squares, each of which can store a single symbol taken from a finite set of possible symbols; and a read-write head that moves along the tape and transmits information to and from the control unit. The machine was instructed by a table that determined the order of operations.

During World War II, Turing worked at the British Foreign Office on assignments that were top secret. For his work, he was awarded the Order of the British Empire (OBE). Some credit him with helping crack the secret German military code.

In 1945, Turing declined a Fellowship at King's College Cambridge to join the Mathematics Division at the National Physical Laboratory (NPL). Earlier work on computability and his expertise in electronics fueled his desire to work on an electronic computer. He called the machine he designed the Automatic Computing Engine (ACE). After Turing left NPL, a pilot model based on his ideas (the Pilot ACE) was built in 1950. It became the successful DEUCE computer.

Elected a fellow of the Royal Society in 1951, Turing tragically committed suicide in 1954 at 41 after being convicted of homosexuality, then a crime under British law.

Excerpts from Turing's Mind article are reprinted with permission from Oxford University Press.

The second annual Loebner Prize Competition will be held in Boston on November 17th, 1992, and official applications must be postmarked by July 31st. Again, the test will be restricted, and $2,000 and the Loebner Prize medal will be awarded to the designers of the best computer entry. Application materials may be obtained by contacting the Cambridge Center for Behavioral Studies.

I suspect that the quality of programs in 1992 will not be substantially better than those tested last year, but after the second contest I predict rapid improvement. The 1991 winner received the equivalent of perhaps a million dollars in free advertising through press coverage of the event; nearly 30 million "impressions" appeared in various media in a matter of days following the contest. Three articles appeared in the *New York Times* alone, including a front page article the day after the contest. I predict that major American corporations will soon see some advantages to this. Budget money will eventually be allocated for developing suitable

---

02:56:11 Judge 2: I wasn't planning to buy anything and I'm not a liar.

02:56:37 Terminal 5: The liar's punishment is that he is not believed even when he tells the truth. I'm going to cry.... Do you say you wasn't planning to buy anything and you're not a liar for some special reason?

02:57:32 Judge 2:

02:57:45 CHANGE OF JUDGES

02:57:47 Judge 8: Hello. I suppose you love Lewis Carroll.

02:58:47 Terminal 5: I used to drive the caddy, but now I often use the mercedes. Do you often joke around? Hello - how's the weather up there today?

02:59:26 Judge 8: I guess I would have to say I don't joke around much. Some friends accuse me of having no sense of humor. Are you saying that I'm exceptionally tall?

02:59:60 Terminal 5: What were we saying about me? Am I right, or what? That's quite interesting.

03:00:28 Judge 8: All I know about you so far is that you have a caddy and a Mercedes so I guess you're rich, or a crook. Aha! When you mentioned weather, I forgot that you might not be in the same place I am. How's your weather?

03:01:34 Terminal 5: How do you know that I am not a dog. Perhaps in your fantasy we know about each other. Are you a caring person?

03:02:15 Judge 8: I am easily moved and easily recover. I am pretty sure that you are not a dog.

03:02:52 Terminal 5: Don't tell a woman she is pretty - tell her there is no other woman like her in the world!.

03:03:24 Judge 8: That would be true about everybody. We are all unique, she said humorlessly.

03:03:53 Terminal 5: Give every man thine ear, but few thy voice. What was I saying....? Really, everybody?

03:04:23 Judge 8: Yes, everybody is unique. Are you actually responding to what I said for a change? Wow!

03: 04: 49 Terminal 5: If this is coffee, please bring me some tea; but if this is tea, please bring me some coffee. You seem quite positive. Are you superstitious?

03:05:32 Judge 8: bye

03: 05: 38 Terminal 5: See you later, judge - have a good time at the other terminals, and don't forget to write.

03:06:02 Judge 8:

03: 06 :19 CHANGE OF JUDGES

03: 06: 21 Judge 5: I am not sure what is whimsical?

03: 07: 29 Terminal 5: Do you think it's absolutely normal to be not sure what is whimsical? Please make me laugh.... Is it because you are not sure what is whimsical that you came to me?

03: 08 :17 Judge 5: probably,

03: 08: 27 Terminal 5: Probably no invention came more easily to man than heaven. Ipso facto nonsensico..... Please go on.

03:08:57 Judge 5: I am in the clouds now.

03: 09 :14 Terminal 5: The best you get is an even break.... How peurile! How long have you been in the clouds now?

03:09:43 Judge 5: When I sat here

03 :10: 05 Terminal 5: It's a wise man who knows when it is time to move on. Woof....wooof....woof.... Why do you ask?

03 :10: 32 Judge 5: Tell me more about it.

03:10:46 Terminal 5: I want to talk about your love life, and I mean right now!. frown..... What does that suggest to you?

03:11:18 Judge 5: Boy! What about yours first, I am shy.

03:11:46 Terminal 5: When you can't do anything else to a boy, you can make him wash his face!. Are you a caring person? Does Chat question interest you?

03:12:27 Judge 5: yes

03:12:36 Terminal 5: My dear judge, just how far did you get in school? You are sure? There is something rotten in vienna!

03:13:08 Judge 5: Not only in Vienna in Washington too.

03:13:24 Terminal 5: Every man who takes office in washington either grows or swells!

03:13:43 Judge 5: I agree, more Chat please.

03:13:57 Terminal 5: I think you said we you asked my how I was doing so i told ya and i'm sorry i make so many typos but please dont' throw them back at me.? We you asked my how I was doing so i told ya and i'm sorry i make so many typos but please dont' throw them back at me.

03:15:09 Judge 5:

03:19:35 INTERACTION ENDS

entries. By year four or five, I suspect the Japanese will get involved. In short, I think the quest for the thinking computer will soon become as intense as the quest for the Holy Grail. The stakes are similar. A program that passes the Turing Test will be worth billions. Just ask it.

I'd like to see the contest expanded at some point to include Turing-like tests of robotics, speech recognition and synthesis, and pattern recognition. In a week-long tournament, computers would compete against people in each domain. The ultimate outcome? Well, did you ever read Asimov's I, Robot series?

I may be overly optimistic about the future of artificial intelligence. Certainly, several of my colleagues, much older and, by definition, much wiser than I, tell me so. But we'll all have fun exploring the possibilities—even if, someday, and for reasons I cannot now imagine, we're forced to conclude that the Turing Test cannot be passed.

Notes

1, 2, 3. Copyright, 1991, Cambridge Center for Behavioral Studies. Reprinted by permission. Note that the quotes from the 1991 Loebner Competition transcripts are included here character-for-character as they occurred during the contest. Typographical errors and transmissions errors are preserved. Complete transcripts of all of the conversations may be obtained by contacting the Center at 617-491-9020, fax: 617-491-1072. A playback program is also available, which will replay the conversations in real-time, exactly as they occurred during the contest.

ROBERT EPSTEIN earned his Ph.D. at Harvard University in 1981. He is the Program Officer in charge of the Loebner Prize Competition for the Cambridge Center for Behavioral Studies and is also the founder and Director Emeritus of the Center. He is the author of more than 60 papers on creativity, problem solving, and other topics, and of the upcoming book, *Cognition, Creativity, and Behavior: Essays on Generativity and Other Topics in the Analysis of Behavior* (Praeger, New York). He serves on the faculties of Boston University, the University of Massachusetts at Amherst, and the University of California, San Diego.

What's Next?

Following the November 8, 1991, contest, The Computer Museum, Boston, plans to create an interactive exhibit for its Smart Machines Gallery based on conversations of the winning computer program and those of a human confederate. The exhibit will dramatize the competition results and educate the public about the importance of the Turing Test and the prospect of intelligent computers.

Each year, the Loebner Prize Competition cash award will be increased. In some years, open-ended Turing Tests will be held. When a computer passes that test, at least $100,000 will be awarded and the prize abolished.

For an overview of the Turing Test, see David L. Waltz, "The Prospects for Building Truly Intelligent Machines," *Daedalus,* 117, Winter, 1988, reprinted in S. Graubard, ed., *The Artificial Intelligence Debate,* 1988, Cambridge, Mass: MIT Press.

The csnet electronic addresses for AAAI are as follows:

AAAI Membership Inquiries:
membership@aaai.org

***AI Magazine* Subscriptions:**
membership@aaai.org

***AI Magazine* Editorial Queries & Letters:**
aimagazine@aaai.org

***AI Magazine* Announcements:**
aimagazine@aaai.org

***AI Magazine* Advertising & Press Releases:**
aimagazine@aaai.org

AAAI National Conference: ncai@aaai.org

AAAI Press:
press@aaai.org

Workshop Information:
workshops@aaai.org

Innovative Applications Conference:
iaai@aaai.org

Spring Symposium Series:
sss@aaai.org

Other Inquiries and Requests:
admin@aaai.org