



Chatterbot Fools Judge Into Thinking It's Human at Competition

BY AARON SAENZ ON OCT 28, 2010 | ARTIFICIAL INTELLIGENCE

5,901 6 ☆



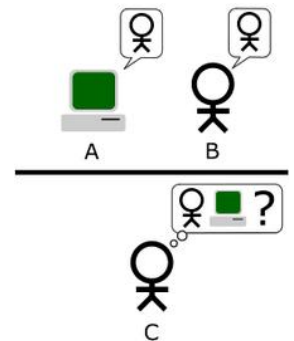
The next time you're on IM, you better watch out, computers are getting better and better at pretending to be human. At the recent [Loebner Prize](#) competition in LA, a chatterbot named [Suzette](#) won first place after convincing a judge that it was a real person. To make matters even more impressive, the judge was none other than the competition's organizer, [Professor Russ Abbot of Cal State](#). Not bad, Suzette, not bad at all. The victory let creator Bruce Wilcox walk off with a \$3000 prize. Before you get too worried about rogue computer programs impersonating the president, however, you should know

that these simple [Turing Tests](#) are a little rigged in the chatbot's favor. It's cool that Suzette fooled Abbot into thinking it was a human, but we're still a long way from AIs entering our society as equals.

The Loebner Prize is awarded each year to the artificial intelligence that can best imitate a human in text-based conversations. These AI chatterbots converse with a judge at the same time as the judge is typing back and forth with a person on a computer. The judge must decide which of the 'people' he's talking with is a human, and which is the chatterbot. It's a pretty standard (low level) interpretation of the famous Turing Test.


One of the big problems with these kinds of tests are that the human participants know they are taking them. Humans can be real bastards sometimes, and it's unclear how often the participants just want the judges to fail. They can talk like a chatbot to throw the judges off the scent. [According to New Scientist](#), this is what Russ Abbot thinks happened, "The human participants were students and two of the judges were professors. Perhaps they simply wanted to fool the judges."

I don't think Abbot's trying to cover his ass here, the results of the competition tend to back up his story. 'Timmy', the human paired with Suzette when she fooled Abbot, tended to make his AI competitors look good. There were four rounds (each 25 minutes long) where each judge got a chance to chat with an AI and a human. Timmy was paired with each of the four AI finalists (a different one each round). For three out of the four judges, the AI




The Loebner Prize is a classic example of the Turing Test, where judges try to determine who is a chatbot and who is a human. But what happens when the humans enjoy fooling the judge?


Latest Popular Editor's Picks



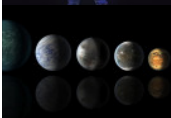
Universities Should Rethink Academic Ideals—Joining Industry Supercharges Research and Tech




12 Industries Disrupted by Tech Companies Expanding Into New Markets




Resurrecting the Dead—Bringing Back the Dearly Departed Digitally




Are We Alone in the Universe? The \$100 Million Search for an Answer



Scientists Connect Brain to a Basic Tablet—Paralyzed Patient Googles With Ease



Wells Fargo Wants to Let You Make Million-Dollar Wire Transactions With Your Face and Voice



Meet the Engineer Bringing Wearable Sensors and AI to Autism Therapy

different one each round). For three out of the four judges, the AI paired with Timmy was ranked most human. It wasn't just Suzette then, but most of the AIs, that seemed more human when compared to Timmy. Results like these make you want to call shenanigans on the whole Turing Test concept.


Even if winning the Loebner Prize isn't a good indication of how human a chatterbot can be, qualifying for the Loebner competition is. Consistently, the finalists at these annual meetings represent the best chatbots you can find. Rollo Carpenter, who took third this year, is the creator of top notch conversationalists like Jaberwacky, Joan, and Cleverbot. The latter regularly fools humans into thinking they are chatting with other people, as you can see in the [comments section](#) or [our story on the AI](#). Richard Wallace, who came in second, is the creator of [Alicebot](#), one of the most famous and successful chatbots in the world. The last finalist this year, Robert Medeksza, sells his own chatbot engine, [UltraHal](#), that you can test out on the web. Along with Suzette, these are all great examples of how AI programs can pull us into conversation.

Keeping us talking, however, isn't the same as fooling us into thinking a program is a human. You can spend hours playing a video game, but you probably (hopefully) don't think any of those characters are real. And we need to remember that the Turing Test (in its traditional incarnations) is a very subjective examination. A chatbot that fools you may not fool me, and vice versa. Likewise, there are humans that seem much more robotic than others. Length of the interaction, and range of topics is also a big factor. Some of the time I have trouble figuring out which of the comments on the Hub come from real people, but I think a five minute conversation would clear up any doubts rather easily.

Instead of focusing on which chatterbots are most human, perhaps we should focus on which are most helpful. Can a bot interact with a lonely person and make them feel better? Can an [AI program answer the phone on a customer service line](#) and not anger you to the point that you want to murder someone? These qualifications are probably much more important than the Turing Test, and they are already taking place in the real world. I congratulate Bruce Wilcox and Suzette for winning the Loebner Prize, but that's just a first step towards where human-simulating artificial intelligences need to be headed. The goal shouldn't be to trick humans into thinking they are talking with a human instead of a computer, the goal should be increasing the quality of the conversation until it doesn't matter either way.

[image credits: CricketSoda, Bilby via WikiCommons]
[sources: [Loebner Prize](#), [New Scientist](#)]


About Latest Posts



Aaron Saenz

RELATED TOPICS: [BRUCE WILCOX](#) [LOEBNER 2010](#) [LOEBNER PRIZE](#) [RICHARD WALLACE](#) [ROBERT MEDEKSZA](#) [ROLLO CARPENTER](#) [RUSS ABBOT](#) [SUZETTE](#) [TURING TEST](#)


Discussion — 6 Responses



Anonymous · 6 years ago

I wanna see google enter their own Chatbot next year. Then they can use it to assist in searches

[Log in to Reply](#)



Brad Arnold · 6 years ago

Just about the silliest and weirdest thing I've ever seen is