Make it Theory workshop
Back in the late '40s, Alan Turing, fresh from his success as part of the team that cracked the Enigma machine, turned his thoughts to machine intelligence. In particular, he considered the question: ‘Can machines think?’ In 1950, he published a paper called Computing Machinery and Intelligence in the journal Mind that summarised his conclusions. This paper became famous for a simple test he'd devised: the Turing Test.

The problem he found in particular was the word 'think' in his original question. What is thinking and how do we recognise it? Can we construct an empirical test that conclusively proves that thinking is going on? He was of the opinion that the term was too ambiguous and came at the problem from another angle by considering a party game called the Imitation Game.

Parlour games
In the Imitation Game, a man and a woman go into separate rooms. The guests don’t know who is in each room. The guests send the subjects written questions and receive answers back (in Turing’s day these were typewritten so that no clues could be found by trying to analyse the handwriting; today we can imagine email or Twitter as the medium). From the series of questions and answers from each subject, the guests try to work out who is the man and which the woman. The subjects try to muddy the waters by pretending to be each other. The fi rst part of Figure 1 shows this game. Turing wondered what would happen if we replaced the man or the woman with some machine intelligence. Would the guests guess the machine versus human more accurately than they would the man versus the woman? Turing’s insight was to rephrase ‘Can machines think?’ into a more general ‘Can machines imitate how humans think?’ This is the original Turing Test, and is shown in the second part of Figure 1.

Over the years, this Turing Test has changed into the simpler test we know today: can we determine whether the entity at the end of a communications link is computer program or human, just by asking questions and receiving answers? It’s is still a variant of the Imitation Game, but now it’s much simpler – possibly too simple.
The first program to try to pass the Turing Test was a program called ELIZA, a program that pretended to be an empathic or non-directional psychotherapist, written in the period 1964–1966. ELIZA essentially parsed natural language keywords and then, through pattern matching, used those keywords as input to a script. The script (and the database of keywords and the subject domain) was fairly small, but nevertheless ELIZA managed to fool some people who used it. The reason for using a psychotherapist as the domain is that it lends itself to being able to respond to statements with questions that repeat the statement (“I like the colour blue.” “Why do you like the colour blue?”) without immediately alerting the human that the therapist on the other end has no applicable domain. ELIZA essentially parsed language and to identify keywords that can then be used to further the conversation, either in conjunction with an internal database of facts and information, or with the ability to use those keywords in scripted responses. These responses give the illusion to the human counterpart that the conversation is moving forward meaningfully. 

ELIZA was so successful that its techniques were used to improve the interaction in various computer games, especially early ones where the interface was through typed commands.

**Chatterbots**

Despite its simple nature, ELIZA formed the basis of a grand procession of programs designed and written to try to pass the Turing Test. The next such was known as PARRY (written in 1972, and designed to be a paranoid schizophrenic), and they spawned a whole series of more and more sophisticated conversational programs called chatterbots. These programs use the same techniques as ELIZA to parse language and to identify keywords that can then be used to further the conversation, either in conjunction with an internal database of facts and information, or with the ability to use those keywords in scripted responses. These responses give the illusion to the human counterpart that the conversation is moving forward meaningfully.

“Chatterbots are still too simple for us to be fooled for long. They don’t have the depth of day-to-day knowledge that makes conversations interesting in the real world, and any conversation over a few minutes reveals that paucity of information.”

A more recent development was the episode of the quiz show Jeopardy! in which two human contestants played against IBM’s Watson computer (although given the fact that Watson has 2,880 processor cores and 16 terabytes of RAM, perhaps ‘computer’ is too simple a term). Watson was not only programmed with a natural language interface that could parse and understand the quiz questions, but also had a four-terabyte automated online ‘customer service rep’ it calls Sarah. Using the chat program is quite uncanny. As you can see in Figure 2, the answer to my question (“How do I see the payments I made in January”) appeared instantly and the natural language evaluation processing is extremely efficient. Notice that the word ‘payment’ is recognised and the more accurate ‘transaction’ is used in the reply.

Such automated online assistants are available 24/7, and are helping reduce the loads on normal call centres. Aetna, a health insurer in the US, estimates that ‘Ann’, the automated assistant for its website, has reduced calls to the tech support help desk by 29 per cent. Of course, there are downsides to chatterbots as well. It’s fairly easy to write chatterbots to spam or advertise in chat rooms while pretending to be human participants. Worse still are those that attempt to cajole their human counterparts into revealing personal information, like account numbers.

**Depth of knowledge**

In reality though, chatterbots are still too simple for us to be fooled for long. They don’t have the depth of day-to-day knowledge that makes conversations interesting in the real world, and any conversation over a few minutes reveals that paucity of information.

A more recent development was the episode of the quiz show Jeopardy! in which two human contestants played against IBM’s Watson computer (although given the fact that Watson has 2,880 processor cores and 16 terabytes of RAM, perhaps ‘computer’ is too simple a term). Watson was not only programmed with a natural language interface that could parse and understand the quiz questions, but also had a four-terabyte automated online ‘customer service rep’ it calls Sarah. Using the chat program is quite uncanny. As you can see in Figure 2, the answer to my question (“How do I see the payments I made in January”) appeared instantly and the natural language evaluation processing is extremely efficient. Notice that the word ‘payment’ is recognised and the more accurate ‘transaction’ is used in the reply.

Such automated online assistants are available 24/7, and are helping reduce the loads on normal call centres. Aetna, a health insurer in the US, estimates that ‘Ann’, the automated assistant for its website, has reduced calls to the tech support help desk by 29 per cent. Of course, there are downsides to chatterbots as well. It’s fairly easy to write chatterbots to spam or advertise in chat rooms while pretending to be human participants. Worse still are those that attempt to cajole their human counterparts into revealing personal information, like account numbers.

**Depth of knowledge**

In reality though, chatterbots are still too simple for us to be fooled for long. They don’t have the depth of day-to-day knowledge that makes conversations interesting in the real world, and any conversation over a few minutes reveals that paucity of information.

A more recent development was the episode of the quiz show Jeopardy! in which two human contestants played against IBM’s Watson computer (although given the fact that Watson has 2,880 processor cores and 16 terabytes of RAM, perhaps ‘computer’ is too simple a term). Watson was not only programmed with a natural language interface that could parse and understand the quiz questions, but also had a four-terabyte automated online ‘customer service rep’ it calls Sarah. Using the chat program is quite uncanny. As you can see in Figure 2, the answer to my question (“How do I see the payments I made in January”) appeared instantly and the natural language evaluation processing is extremely efficient. Notice that the word ‘payment’ is recognised and the more accurate ‘transaction’ is used in the reply.

Such automated online assistants are available 24/7, and are helping reduce the loads on normal call centres. Aetna, a health insurer in the US, estimates that ‘Ann’, the automated assistant for its website, has reduced calls to the tech support help desk by 29 per cent. Of course, there are downsides to chatterbots as well. It’s fairly easy to write chatterbots to spam or advertise in chat rooms while pretending to be human participants. Worse still are those that attempt to cajole their human counterparts into revealing personal information, like account numbers.

**Depth of knowledge**

In reality though, chatterbots are still too simple for us to be fooled for long. They don’t have the depth of day-to-day knowledge that makes conversations interesting in the real world, and any conversation over a few minutes reveals that paucity of information.

A more recent development was the episode of the quiz show Jeopardy! in which two human contestants played against IBM’s Watson computer (although given the fact that Watson has 2,880 processor cores and 16 terabytes of RAM, perhaps ‘computer’ is too simple a term). Watson was not only programmed with a natural language interface that could parse and understand the quiz questions, but also had a four-terabyte automated online ‘customer service rep’ it calls Sarah. Using the chat program is quite uncanny. As you can see in Figure 2, the answer to my question (“How do I see the payments I made in January”) appeared instantly and the natural language evaluation processing is extremely efficient. Notice that the word ‘payment’ is recognised and the more accurate ‘transaction’ is used in the reply.

Such automated online assistants are available 24/7, and are helping reduce the loads on normal call centres. Aetna, a health insurer in the US, estimates that ‘Ann’, the automated assistant for its website, has reduced calls to the tech support help desk by 29 per cent. Of course, there are downsides to chatterbots as well. It’s fairly easy to write chatterbots to spam or advertise in chat rooms while pretending to be human participants. Worse still are those that attempt to cajole their human counterparts into revealing personal information, like account numbers.

**Depth of knowledge**

In reality though, chatterbots are still too simple for us to be fooled for long. They don’t have the depth of day-to-day knowledge that makes conversations interesting in the real world, and any conversation over a few minutes reveals that paucity of information.

A more recent development was the episode of the quiz show Jeopardy! in which two human contestants played against IBM’s Watson computer (although given the fact that Watson has 2,880 processor cores and 16 terabytes of RAM, perhaps ‘computer’ is too simple a term). Watson was not only programmed with a natural language interface that could parse and understand the quiz questions, but also had a four-terabyte automated online ‘customer service rep’ it calls Sarah. Using the chat program is quite uncanny. As you can see in Figure 2, the answer to my question (“How do I see the payments I made in January”) appeared instantly and the natural language evaluation processing is extremely efficient. Notice that the word ‘payment’ is recognised and the more accurate ‘transaction’ is used in the reply.

Such automated online assistants are available 24/7, and are helping reduce the loads on normal call centres. Aetna, a health insurer in the US, estimates that ‘Ann’, the automated assistant for its website, has reduced calls to the tech support help desk by 29 per cent. Of course, there are downsides to chatterbots as well. It’s fairly easy to write chatterbots to spam or advertise in chat rooms while pretending to be human participants. Worse still are those that attempt to cajole their human counterparts into revealing personal information, like account numbers.

**Depth of knowledge**

In reality though, chatterbots are still too simple for us to be fooled for long. They don’t have the depth of day-to-day knowledge that makes conversations interesting in the real world, and any conversation over a few minutes reveals that paucity of information.

A more recent development was the episode of the quiz show Jeopardy! in which two human contestants played against IBM’s Watson computer (although given the fact that Watson has 2,880 processor cores and 16 terabytes of RAM, perhaps ‘computer’ is too simple a term). Watson was not only programmed with a natural language interface that could parse and understand the quiz questions, but also had a four-terabyte automated online ‘customer service rep’ it calls Sarah. Using the chat program is quite uncanny. As you can see in Figure 2, the answer to my question (“How do I see the payments I made in January”) appeared instantly and the natural language evaluation processing is extremely efficient. Notice that the word ‘payment’ is recognised and the more accurate ‘transaction’ is used in the reply.

Such automated online assistants are available 24/7, and are helping reduce the loads on normal call centres. Aetna, a health insurer in the US, estimates that ‘Ann’, the automated assistant for its website, has reduced calls to the tech support help desk by 29 per cent. Of course, there are downsides to chatterbots as well. It’s fairly easy to write chatterbots to spam or advertise in chat rooms while pretending to be human participants. Worse still are those that attempt to cajole their human counterparts into revealing personal information, like account numbers.

**Depth of knowledge**

In reality though, chatterbots are still too simple for us to be fooled for long. They don’t have the depth of day-to-day knowledge that makes conversations interesting in the real world, and any conversation over a few minutes reveals that paucity of information.

A more recent development was the episode of the quiz show Jeopardy! in which two human contestants played against IBM’s Watson computer (although given the fact that Watson has 2,880 processor cores and 16 terabytes of RAM, perhaps ‘computer’ is too simple a term). Watson was not only programmed with a natural language interface that could parse and understand the quiz questions, but also had a four-terabyte automated online ‘customer service rep’ it calls Sarah. Using the chat program is quite uncanny. As you can see in Figure 2, the answer to my question (“How do I see the payments I made in January”) appeared instantly and the natural language evaluation processing is extremely efficient. Notice that the word ‘payment’ is recognised and the more accurate ‘transaction’ is used in the reply.

Such automated online assistants are available 24/7, and are helping reduce the loads on normal call centres. Aetna, a health insurer in the US, estimates that ‘Ann’, the automated assistant for its website, has reduced calls to the tech support help desk by 29 per cent. Of course, there are downsides to chatterbots as well. It’s fairly easy to write chatterbots to spam or advertise in chat rooms while pretending to be human participants. Worse still are those that attempt to cajole their human counterparts into revealing personal information, like account numbers.
database of structured and unstructured information from encyclopedias (including Wikipedia), dictionaries, thesauruses, and other ontologies. In essence, Watson has ‘knowledge’ about many things and the software to query that knowledge, to form hypotheses about that knowledge, and to apply those hypotheses to the questions posed.

Parsing natural language

Although Watson may be seen as the current best contender for passing the Turing Test, there are still huge issues with the natural language into which it was put, perhaps the hardest part of the software to write. One of the questions asked (or rather, given the nature of the quiz show, one of the answers for which the contestants had to provide the question) was: “Its largest airport was named for a World War II hero; its second largest, for a World War II battle”, and Watson failed to parse the sentence properly, especially the part after the semicolon, causing it to reply with “What is Toronto?” when the answer should have been “What is Chicago?”.

Natural language can be parsed in two ways: either through a strict semantic analysis and continually improving that analysis, or through a looser ‘last in, first out’ philosophy of analysis. In essence: using a database of hundreds of millions of words, combined with interesting algorithms and statistical methods to improve the result. This kind of algorithm is used by Google Translate: by using a huge corpus of texts and translations by human linguists, the quality of on-demand translations can be improved greatly. Google Translate uses the power of the crowd to translate text rather than strict language models to such an extent that the resulting translations can be improved greatly. Google Translate is one such neural network: Viola and Jones used a database of thousands of faces from the internet to tune their network to recognise faces accurately (according to their paper, they used 4,916 images of faces and 9,500 images of non-faces that they sliced and diced in various ways). The training took weeks.

One of the problems with the Turing Test is that it invites programs with ever more complex conversation routines that are just designed to fool a human counterpart. Although researchers are investigating how to create a ‘strong AI’ that can converse with, and possibly understand, a human, those hypotheses are still speculative. The best contender for passing the Turing Test, or at least being considered AI, is a program named Watson.

Neural networks

Many specific-AI systems use a neural network as the basis of the software. This is a software encapsulation of a few neurons, also emulated in hardware and known as perceptrons. Just like our neurons, perceptrons receive stimuli in the form of input signals, and fire off a single signal as output, provided the sum of (or the mix of) input signals is greater than some value. Neural networks need to be trained. The training process involves running examples of features in the problem space the neural network, observing the outputs, comparing them with the desired outputs, then tweaking the configuration of the perceptrons to make the output closer to the expected results.

As you’ve seen, we can view AI through two incompatible lenses. The first is what we’ve grown up with: a computer system and software that can emulate a human being to the extent that it can fool judges using the Turing Test. It’s known as strong AI and is the subject of many science fiction movies, the most famous being HAL 9000 from 2001: A Space Odyssey. The second is perhaps more interesting, because it affects us in our daily lives: specific AI that solves single problems, and that would have been the subject of SF novels just a few years ago. What new specific AI awaits us in the next five years? PGP

Julian M Bucknall has worked for companies ranging from TurboPower to Microsoft and is now CTO for Developer Express.

feedback@peplus.co.uk

---

**Figure 2:** Sarah, the PayPal automated online assistant, is an example of a chatterbot.