Irish Communication Review

Volume 16 | Issue 1 Article 11

June 2018

Is Google Self-aware?

Gerry Heapes
Technological University Dublin, gerard.heapes@tudublin.ie

Is Google self-aware?

Gerry Heapes

Abstract

The field of artificial intelligence evolves incrementally with gradual improvement over time and has relied on the Turing Test as a measure of progress. However human standards of intelligence measurement may not be appropriate to current developments. The platform of cloud computing now provides a means of implementing a kind of ubiquitous awareness unknown to humans before now and a means of augmenting human intelligence. The level of awareness held by Google is explored and some recent developments in the uses of AI programmes for social media are covered. The misinterpretation of these developments is explored and a solution proposed. Finally, a summary of future developments in AI is presented.

The emergence of artificial intelligence (AI)

If the origin of intelligence is examined, particularly artificial intelligence (AI), the complexity of the subject soon becomes apparent. AI has been studied for as long as practical computers have existed and yet has proven to be an elusive prospect for computer scientists. The task of building an intelligent machine comparable in ability to a human mind has turned out to be much more difficult than expected largely due to the fact that the scope of intelligence is almost unlimited and its parameters are difficult to define. AI ranges from simple machines using search algorithms to play board games up to modern neural networks using enormous processing power and self-teaching techniques. The

emergence of large scale and viable AI has serious implications for every aspect of the way in which computers are used by modern society.

The critical importance of finding a means of harnessing and controlling knowledge and information has long been a concern of academics and even more so since the end of World War II when computers became a viable field of research (Bush, 1945).

The current view of the emergence and evolution of human and animal intelligence follows a Darwinian model in which small incremental mutations are passed on from generation to generation as improvements which tend to make the species more likely to propagate further. The process is normally gradual but is occasionally punctuated by rapid bursts of significant improvement in which some emergent characteristic is obvious, such as improved eyesight.

The point to note is that the eye as we know it did not emerge fully formed overnight but arose gradually as an improvement in our ability to perceive light, evolving through many less complex but nevertheless useful iterations. The question is posed “Of what use is half an eye?” The answer is little use, because it is probably non-functional, but if the question is rephrased as “Of what use is 50% vision?” the answer is that it would almost certainly save your life and possibly help find you a mate with whom to pass on your genes.

The evolution of animal intelligence, including our own, has followed a similar path and the growing emergence of AI might reasonably be expected to follow the same pattern. Assuming that it does, we might expect that a gradation of artificial intelligence has already emerged, and will continue to develop with a corresponding gradation of usefulness and independence. The evidence for this emergence may not be obvious, but just because a machine cannot currently solve great mathematical, scientific or cultural problems does not mean it is useless.

The assertion that machines cannot possibly think in the same way as humans is meaningless because machines can now access, store, process and make inferences from quantities and types of data that no human possibly could. It follows then that the standards and tests for human or animal intelligence may not be appropriate or even meaningful when applied to artificial intelligence.

The Turing Test

In 1936, the English mathematician Alan Turing published a fundamentally important paper entitled "On Computable Numbers with an Application to the Entswcheidungsproblem", (Turing, 1936**)** which essentially laid the theoretical foundations for what a computer might be capable of and even what was computable at all.

In a second important paper published in 1950, Turing proposed a method to decide whether machines can think, commonly known as the Turing Test (Turing 1950), long before the implications of artificial intelligence would become apparent. The test, referred to informally as 'The Imitation Game', is essentially simple, proposing that the existence of a computer whose responses to questions are indistinguishable from those of an intelligent human would demonstrate that machines can think.

It may be crudely stated as: 'If the machine answers questions from a suspicious human sufficiently well to persuade them into thinking that it is human then we must conclude that the machine is intelligent.'

This simple test has very far reaching consequences because it defines a general but clear end goal rather than solving a specific problem such as teaching a machine to play chess autonomously. The generality of the test requirements has driven research on a wide range of AI related questions like natural language processing – learning in real time from conversations, communicating naturally and trying to grasp common sense notions.

The original form of the Imitation Game requires a man, a woman and an interrogator who wishes to determine if the other participants are men or women based solely on their anonymous responses to questions.

The Turing Test replaces one of the participants by a machine and the aim of the interrogator is to deduce which is a human and which is a machine based solely on their anonymous responses to questions by the interrogator. The interrogator knows that there is one human and one machine so the test asks the computer to persuade a suspicious human.

For Turing, such a test was sufficient proof of machine intelligence and the paper continues with an introduction to digital computers and their use for arbitrary

computation, asking whether the appearance of intelligent or sentient behaviour is a computable problem at all, or is simply a true but untestable proposition.

Many objections have been raised to the Turing Test because it could be passed by a machine incapable of writing poetry or music with their associated emotional responses. Turing himself argues that an observer cannot tell if a machine feels unless they themselves are the machine, making the claim untestable. The machine and program subjected to the Turing test may include elements such as convincing the interrogator that it does in fact feel emotion, even in the absence of proof of that response in a similar way to a human.

This would, in fact, mirror the way in which humans communicate to convince each other of their sentiments without having any real proof of them. The fact is that humans often lie to each other, so an observer might reasonably ask 'Does lying indicate real intelligence?' – a question that has re-emerged following recent developments in AI.

The Turing Test is applicable to only a narrow range of daily human experience, so it may be asked whether it might be possible to simulate emotional responses, strong motivations or problem solving skills similar to, but not identical to, those of a human observer. If a machine were to pass the Turing Test, does it prove intelligence or is there much more to consider before judgement is passed? It must also be considered that a young child subjected to the Turing Test might fail it for lack of life experience or communication skills but not for lack of the ability to think. Thus it may be possible for a computer to 'think' but still not pass the Turing Test.

AI platforms

Currently, the best hope for AI platforms are devices known as Deep Neural Networks (DNN), where input data is filtered through layers of artificial neurons mimicking the architecture of the brain and extracting recognisable patterns in images for machine vision, and text or sound for natural language processing. The processing is achieved in real time by using massively parallel processors with millions of processing units and also by exploiting the parallel nature of much of the input data itself. They form the base platforms for AI applications like Facebook Deep Text and Google DeepMind; the latter has recently beaten

the world's best players at the Chinese board game of Go, a game considered much more difficult than chess (GDM 2017).

The normal limits of parallel processing (Amdahl's Law) do not apply because the problem size is scaled up to suit whatever available processing power is presented (Gustafson's Law), which currently is practically unlimited.

The use of deep learning techniques has changed the way in which natural language is processed, and the emphasis now is on extracting semantic relations which are applicable across many languages and not simply the assignment of a database code for each specific word in a given translation. Facebook Deep Text is independent of the written or spoken language to which it is applied (FDT 2017).

Google DeepMind's aim is to produce a set of general purpose algorithms linked to provide a self-teaching AI system called an agent. This approach is quite different from the traditional application of AI to solve very specific problems using expert systems which encoded a particular knowledge base and were not flexible to learning outside these parameters.

In image recognition, the level of accuracy achieved by DeepMind is comparable with a human in about 95% of cases and it forms the basis for Google+ Image Search. The system will accept a word search and retrieve images based upon this request even when the images are unlabelled or have no distinguishing tags or data attached. It is also used to accelerate speech recognition using the same generic deep learning algorithms and architectures. A set of similar networks are in use for fraud detection, handwriting recognition and translation replacing older rule based AI with DNN technologies. The question arises of how do you measure the IQ of an AI system? Human standards may not be appropriate and the tendency toward anthropomorphism must be avoided (Liu 2016).

Cloud computing

All of the aforementioned AI tools are facilitated and enabled by the use of 'cloud computing' architectures, in which the processing power of many thousands of computers may be brought to bear on a single given task flexibly and on demand. Currently, most internet applications run upon cloud based platforms and the physical infrastructure is housed in large server farms connected by very high speed data links.

The basic premise of the cloud is to turn computing power into a billable utility charged for on demand like gas or electricity. It makes available to a mass market a huge range of computing services and processing power enabling businesses to affordably outsource their computing infrastructure but it also enables the operators to run and maintain virtually unlimited processing ability at very low cost. It also gives the cloud operators access to and control over all the data passing through their systems along with the capacity to acquire, store and process an unlimited quantity of data.

The data gathered is then mined using the cloud itself with AI algorithms to extract useful information for commercial gain. Mass market users pay for free services by giving up a measure of their privacy through their data. The gathering of data takes place from millions of input points simultaneously and nothing is discarded, with the permanent storage capacity effectively infinite and available round the clock. This allows for a type of 'ubiquitous awareness' previously unavailable and of such a depth and extent that large scale cloud applications using AI have enabled a type of 'augmented human intelligence' unprecedented in human history.

Cloud based AI systems of this type, then, have the ability to take in and store unlimited quantities of live data, use DNN processors and algorithms to adapt in real time, learn from their environment and enable them to react to it in a more intelligent way.

The systems do not only react but also act independently with foresight anticipating many possible different scenarios at any given moment and following due consideration make choices to reach an optimal outcome. The ability to react to changing circumstances and to plan for possibly unforeseen outcomes are both key markers for human intelligence. Finally, systems of this type have the unprecedented ability to interact in real time with literally millions of individuals at once and also to facilitate and concurrently monitor communications between those individuals, an ability unknown up to now.

Is Google self-aware ?

If cloud based applications like Google are 'ubiquitously aware', are they self-aware? Would they, for example pass the Turing Test as described? A variation of the Turing Test has been proposed by Nicholas Negroponte, a co-founder of

the MIT Media Lab, and is described in a book by Stewart Brand (1988). In this modified Turing Test an AI machine is judged on its ability to work in partnership with humans rather than facing suspicious interrogation by them. The question proposed is whether or not the machine helps the human subject to reach a specific goal in a manner similar to how another human might.

Such a test would be potentially much more difficult because the machine now requires an even greater ability for natural language processing to allow for human fallibility in the framing of suitable questions and also spotting human errors in communicating those questions to it. Such a redefinition of the test would follow on directly from the already mentioned facility for augmenting human intelligence.

Within this context then the answer to the question 'Is Google self-aware?' is 'probably not', at present, especially in the strict interpretation offered by the Turing Test, but this position is increasingly under threat. To some extent, this is perhaps asking the wrong question, because Google demonstrably does act intelligently and it is increasingly self-directed.

In some circumstances Google probably acts more intelligently than many humans and it definitively does have a type of awareness of and insight into the world which is totally alien to humans.

Such a situation allows, for the first time, an augmentation of human intelligence unlike any partnership in human history. The presence or absence of self-awareness may be currently untestable for AI systems like Google but awareness of the external world is definitely present and should be readily acknowledged by society at large.

Google knowledge base

The question then arises as to precisely what does Google 'know' about us? The answer to this question may be framed within the range of services available to the casual user of Google.

The core search engine gathers comprehensive details of the subjects and content of what users seek as well as tracking where and when the content was downloaded for both text documents and images. Using AI capabilities it is possible to search images very effectively by using word searches by subject or

physical objects contained within the images. Google Maps tracks the current GPS location of the user to within 1 metre anywhere on the surface of the earth and also keeps a record of any journeys by the user indicated on it. Google Drive offers the user the possibility of storing personal files and Gmail compounds this data source by hosting and monitoring any email transactions which occur between the user and a correspondent. The facility offered by Google Scholar captures a complete record of searches pertaining to a very wide range of academic research streams allowing an overview of scholarly activity on an unprecedented scale. Google Translate removes almost any language barrier left between the user and the world and analysis of translation data could prove to be a rich source of innovation or be used as a marketing tool. Finally the desire to track location for any user may be unnecessary if the user provides a projected plan of their whereabouts using Google Calendars.

In summary Google knows what we seek, our whereabouts past and present, what we store and who we correspond with, what we research, what we translate and what our planned schedules are.

The accessibility and ease of use for all of these services is currently being enhanced by the introduction of new services like voice operation using Google + but even without such additions it is clear that the amount of data gathered and the scale of collection is unprecedented in human history and quite beyond even the most Orwellian outlook.

The ability to gather, hold, integrate and manipulate data on this scale could only be described as intelligent, despite its failure of the Turing Test. The systems operated by companies like Google and Facebook are already omnipresent and are as close to omniscience as has ever existed representing an entity for which normal standards are not adequate.

The Turing Test does not fully take into account the developments of how computing in general and AI in particular have evolved. The idea of an augmented human intelligence presents us with interesting questions about how we interact with it and who or what will control its use. How concerned should we be with the possibility of machine intelligence surpassing our own or, more precisely, surpassing our ability to control it? The likelihood is that machine abilities will eventually grow beyond us so how will we react if and when they do? A recent incident may give a glimpse into this question.

A tale of two bots

A chatbot is an AI driven program used as a means of interfacing with web applications to make them more accessible, user friendly and more engaging by acting like a human, rather than offering a text based hierarchical set of choices to the user. Facebook AI Research recently published the outcome of a set of experiments it ran concerning their use in negotiation scenarios and how surprisingly successful they were at imitating human responses.

The researchers ran a series of autonomous experiments to allow two chatbots to negotiate with each other following training and exposure to large amounts of real data from human negotiation techniques (FDT2, 2017).

One outcome was that the bots evolved their own syntax for communication independently from the programmed model once their training was complete and they were left to run unattended. The syntax was a kind of restructured English similar to the code used to run the bots but whose precise meaning was unknown to the developers and not programmed in by them.

The experiment was terminated as a result of the inability of the designers to fully understand what communication was taking place between the two chatbots. This is commonly encountered in AI research and there has already been at least one recent similar instance at Microsoft AI research, also involving chatbots.

The reasoning of the experimenters was simply that no further useful data could be gathered by them within the parameters applied and hence the experiment was terminated at that point. The results were formally published but were widely misinterpreted. News outlets irresponsibly accused Facebook of panic and worse in the wake of this incident, implying that the machines had taken over in some way and had to be killed off. The resulting stories were sensationalist at best and selectively chose one aspect of the work to the exclusion of all else.

Among the factual findings published in the research were:

1. Goal based AI models negotiate harder than humans.
2. AI models produce novel meaningful sentences of their own.
3. Multi-sentence coherence is still a problem for chatbots.
4. Models taught themselves how to lie to gain advantage while negotiating, an emergent property which had not been explicitly designed into their programming.

This last finding, the emergence of lying, is probably more significant than the novel syntax encountered because it displays a very human characteristic, but was ignored in the ensuing controversy.

The question that researchers were left with was whether the behaviour was truly emergent or was it reflecting a learned behaviour implicit within the original human negotiation data used to train the AI systems involved.

The potential for the spread of misinformation in the field of AI research is very great and the public perception of the work is also important so the necessity for informed judgement should be acknowledged by all.

The ability of machines to imitate humans is improving to a point where concern is natural and a transparent dialogue is absolutely necessary between researchers and responsible news outlets to clarify facts and dispel fears.

Conclusions

There is a gradation of ability and usefulness to artificial intelligence which is not immediately apparent and which is becoming pervasive as it evolves incrementally. The standards of human intelligence push an observer to anthropomorphise artificial intelligence by projecting on to it human characteristics in order to make comparison with human capability easier. Such an approach may not be appropriate and is biased at best.

The Turing Test has defined the framework of proof for AI for more than 60 years but is not generally applicable and current developments in computing indicate that it may be possible for a machine to fail the test but still be considered as a thinking entity.

The narrow application of AI has produced machines capable of beating the best human experts at a range of activities most notably board games such as chess and Go and with the emergence of new computing architectures for AI this is set to continue. The thrust of research is now aimed at designing machines which will be self-teaching following training making them independent of human intervention to evolve consequently requiring a new standard of intelligence measurement techniques.

The pervasiveness and processing power of cloud computing platforms using AI algorithms has turned the acquisition and mining of data into the basis for large commercial monopolies to exploit and influence society on an unprecedented scale. This provides operators with a very novel kind of ubiquitous awareness of the world and the possibility of augmenting human intelligence in its endeavours. The resulting AI driven systems display analytic abilities, foresight and planning, decision making based upon input from millions of sources simultaneously and the ability to monitor the actions and gauge the sentiment of large populations in real time.

The AI system represented by Google would currently be unable to pass the strict Turing Test but is evolving to a point where it may soon be able to pass the more difficult and practical requirements of augmenting human intelligence to enable the solution of problems intractable to humans alone.

The presence or absence of self-awareness is substituted at present by an awareness of the external world which is beyond any human ability in terms of its scale and intent. Google AI has a multi-faceted awareness of the real world alien to human experience but potentially open to dangerous exploitation.

The quantity, depth and extent of data gathered by Google enables its operators to form a profile of its users and their behaviour which would have been unimaginable a decade ago and the AI driven systems which are now emerging will further mine these data assets in unforeseen ways for decades to come.

Such systems and their operators represent a new type of augmented human intelligence highlighting ethical issues about the way in which this development might be applied to problem solving in the world as a whole. To some extent AI has surpassed human ability in narrow areas but it will almost inevitably do so in a general sense at some time in the near future possibly as early as the year 2040.

The ability of AI driven machines to directly imitate human behaviour is improving at a surprising rate and it must be noted that if an artificial general intelligence (AGI) emerges with abilities comparable to that of an intelligent human then it will be able to learn at an exponential rate, much faster than a human. Once this is achieved it may be beyond the ability of human operators to control the outcome or subsequent direction of developments potentially

threatening human existence, even inadvertently. AI self-awareness, in the sense of human level consciousness, may be irrelevant because it will be so different to its human counterpart that it may not even recognise human observers as sentient.

An open, frank and public but most of all informed discussion of this is an absolute requirement to dispel fear and fanciful notions which may arise. The role of journalism and responsible media outlets will be critical to the direction and effectiveness of this discussion and the influence afforded via digital media necessitates a more critical and technically informed approach to AI themed research.

In the near term the possibility exists to use current AI augmented intelligence to avoid adverse outcomes, possibly disadvantageous to biological life.

References

Brand, S. (1988) The Media Lab : Inventing the Future at M.I.T. ISBN-13: 978-0140097016 Penguin Books New York, NY, USA ©1988

Bush, V. 1945. As We May Think. *The Atlantic Monthly*. July 1945.

(FDT1 2017) (arXiv:1502.01710)

(FDT2 2017) (arxiv.org/abs/1706.05125v11)

(GDM 2017) (doi:10.1038/nature16961)

(Liu 2016) A Study on Artificial Intelligence IQ and Standard Intelligent Model (https://arxiv.org/ftp/arxiv/papers/1512/1512.00977.pdf)

Turing, A. (1936) ”On Computable Numbers with an Application to the Entswcheidungsproblem”

(Turing 1950)A. M. Turing (1950) Computing Machinery and Intelligence. *Mind 49:* 433-460.