Notes on the topic "Technology and Robotics"

Before I make a start on my notes, I should say something about the title of your Philosophical Forum topic[1].

The title seems to me rather puzzling and, moreover, in a way that reflects a general puzzle or confusion in the cultural ether at large. For by coupling the terms "Technology" and "Robotics" like this, the title suggests robotics is not itself a technology; almost as if robots had an independent existence and robotics was something like the study of robots as a natural phenomenon. One might feel much the same sense of conceptual disquiet if one came across the title "Technology and Electronics" or "Technology and Automotive Engineering".

You wrote:

> What is odd about this particular Forum announcement is that there's been no input from Forum-ers re questions they think of in relation to the topic they choose. I usually receive questions from a few of them, and in the past, there's even been five-minute introductions to the topic of discussion by at least two people at the Forum meeting itself.

I think it is likely that the unusual lack of input from the 'Forum-ers" on this topic has to do with this culturally induced perplexity, which is itself a natural response to various contradictions and mystifications that have surrounded the idea of artificial intelligence (AI) and robotics over the last 60 odd years (i.e. since the arrival of computers).

Three basic distinctions (frequently ignored, or run together, or made the objects of equivocation):

- Weak AI vs. strong AI – a distinction hinging on two quite distinct senses of 'intelligence'
- Robots as figures in fantasy and mythology vs. robots as technology (robotics)
- Robot/computer speech as simulacrum (chatbots) vs. robot/computer speech as natural language mastery

In what follows I treat the concepts of artificial intelligence and intelligent robots as interchangeable. The mathematician Alan Turing is generally credited with giving the notion of artificial intelligence scientific and philosophical respectability in an article published in the philosophical journal Mind ("Computing Machinery and Intelligence", 1950). Turing doesn't mention robots in the article and bases his argument for the possibility of a thinking machine (he doesn't use the term 'artificial intelligence') entirely on computer technology, then in its infancy. From the point of view of machine intelligence, we can treat the two as the same because whatever intelligence a robot has or is imagined to have comes from the computer inside it. As I said in my email, a robot is just a computer with a mechanical body.

[1] These notes were written in July 2015 for MSJ, an emeritus professor of philosophy. MSJ organizes a Philosophical Forum in her hometown. A week or so earlier she had emailed me that "the topic of our next Forum (topics are chosen by Forum-ers) is on technology and robotics".

It also worth pointing out that ‘computer technology’ can be understood in two ways:

- as software, the collection of programs that enable a computer to do whatever it does; e.g. ‘application programs’ such as wordprocessing, spreadsheet, drawing or graphics programs, internet browsers, etc., and, at a lower level, the operating system software which manages and supports application programs. The most widely used operating system for personal computers is Windows, while other types of personal computers like the Macintosh have their own operating systems.
- as hardware, the electronic machinery, the most important component of which is the computer’s ‘processor’, the device which runs or ‘executes’ the programs.

Of these two aspects of computer technology, by far the most important for AI is computer software. Without software, a computer is inert: it can do nothing. Every development in artificial intelligence has been on the software side. Textbooks on Artificial Intelligence are primarily concerned with software techniques and methods (sometimes loosely referred to as ‘algorithms’) for simulating or reproducing in a limited sphere human activities and skills (e.g. playing chess). Computers have become very much faster and more powerful over the last 60 years, but progress in this direction has made no difference to the development of artificial intelligence. In short, the technology of artificial intelligence is to all intents and purposes software technology.

I emphasise the point – i.e. if there can be such a thing as an intelligent machine, the major part of building it will involve developing programs for it rather than creating faster and more powerful computer hardware – because this is another area in which enthusiasts for AI either blur a crucial difference or ignore it altogether. A very good example of this type of confusion is the idea that computers can write prose or even poetry. In one of his later essays, the great English literary critic F.R. Leavis responded very decisively to this idea (see quote §1 from Leavis at the bottom of this document). Leavis was born in 1895. He knew practically nothing about computers, and yet he was fully justified in protesting when told that “A computer can write a poem”.

The claim that computers can write prose or poetry trades on an equivocation between two senses of ‘write’. In the literal sense (as it were, literally literal), a computer can generate meaningful and grammatically correct text in so far it can be programmed to display the text on screen or print it out. But this, of course, does not mean the computer itself is ‘writing’ in the same sense as people write. Instead it is merely executing a program written by a human being (a programmer) which generates the text. The program itself is not intelligent in any sense of the word ‘intelligent’. How could it be? A program is just a list of instructions written in a programming language.

People outside the world of computers and software are sometimes misled into thinking that if a program can spew out meaningful sentences it must be employing methods or principles similar to those involved when human beings write prose or poetry. Nothing could be further from the truth. I quote myself from an email discussion I had some years ago about the Leavis passage:

> In the figurative or derivative sense, it is true that computers can write poetry – but the very fact that in describing what they do, or rather what the software does, I and other people tend to use words like ‘generate’ or ‘produce’ shows that ‘write’ has a different meaning here to that which it has when applied to human beings. If I

> wrote a program that spliced, mixed and matched, and combined lines from different poems by Alexander Pope to produce new rhyming couplets, neither you nor anyone else, I think, would want to say the program was writing poetry. Now the programs that generate poetry are more sophisticated than this – but not in any significant way, not in a way that could somehow engender intelligence. Essentially they work like this: they manipulate words, perhaps using a rhyming dictionary, perhaps consulting a lexical database of some kind, according to simple grammatical and parsing rules. And, of course, they do so mindlessly. The words are not meaningful to the computer – not by any conceivable philosophical measure of meaning, and not by the standards of the most scientistic, logistic semantic theory. What's more, nobody – except a handful of cranks and fantasists – thinks of such software as an adequate model for human intelligence. It's no more sensible to say a computer can write poetry than it is to say a computer with Microsoft Word's grammar checker can (literally) edit your text, or a thermometer can (literally) take your temperature, or an aeroplane can (literally) fly like a bird. To think otherwise is not to be open-minded about technology or to affirm your faith in what science can achieve – it's crude science fiction, tantamount to a belief in fairy tales and fantasies.

In this case, several equivocations and confusions are in play:

> 1. An equivocation between two senses of 'write': one in which the word refers to the writing human beings do, an activity which requires consciousness, sentience, and a mastery of natural language (i.e. the ability to speak and understand language); the other in the limited and technical sense in which a computer program produces text as output (e.g. to the screen, disk, or printer).
>
> 2. A confusion between a piece of software, a program, and the computer that runs or 'executes' the software. The most we can say about computer-generated text is that it is produced by a program which has been written by a programmer. Only if we ignore the two aspects of computer technology (hardware and software) are we likely to treat the computer as an autonomous, self-contained device – imagined to be equipped in some way with volition and intelligence – and talk about it 'writing' text as if it were engaged in the same activity as people.

I won't pursue the argument further here, but I think one can diagnose and explain the mistaken line of thinking behind the belief that can computers can master language (write, read, speak, and understand). It comes from the universal human tendency to attribute animate qualities to inanimate objects. In persuading ourselves that a computer is writing or speaking or understanding language in the same way that people do (as opposed to giving the impression of doing so or creating a simulacrum of natural language ability), we are succumbing to this extremely powerful tendency. When we attribute intelligence and sentience to a computer we are projecting human capacities onto an inanimate object. In a word, we 'personify' the computer. In so doing we are engaged in exactly the same exercise in fantasy and imagination as the belief in talking animals or talking trees. Of course, in fantasy fiction and film we don't really believe animals or trees can talk. The all-important difference with computers is that because they can be programmed to give the impression of reacting to, or interacting with, people, it is very much easier to take the idea of a genuine non-biological intelligence (a machine intelligence) seriously. Indeed, when the computer is lodged in a mechanical body (a robot) and to some extent

behaves in a life-like fashion (albeit robotic and mechanical in its movements, lacking any natural animal fluidity), it is almost impossible not to give in to the animistic tendency.

The objections commonly raised to the diagnosis and argument above go like this:

- You're talking about early experiments and research efforts in artificial intelligence carried out nearly 50 years ago. Computer technology, both on the hardware and software side, has made enormous progress since then. In particular, the academic field of artificial intelligence and industrial applications of AI, which barely existed when Leavis was writing, are now well-established and flourishing. Look at the achievements of AI in the areas of speech recognition, machine translation, expert systems, and automated knowledge bases. Look at such triumphs of computer intelligence as IBM's Watson system, which took on and beat human contestants in the TV show Jeopardy. There are many other instances of AI in action, not least among them applications in computer games, chess playing (IBM's Deep Blue machine beat the world champion Kasparov at chess), sophisticated pattern recognition (e.g. the ability to recognize faces), Google's intelligent search capabilities, data mining, and Honda's humanoid robots.

- In any case, by choosing the example of a computer writing poetry you've rigged the argument in your favour. Most people in the AI field don't seriously think computers will be able write poetry. That's something that calls for peculiarly human qualities like creativity and emotion. What AIs will be able to do, and where they will excel, is employ a very high level of intelligence, moreover an intelligence unsullied by such human frailties as emotion, feelings, bias, superstition, ideology, and other irrational elements in the way people think. An artificial intelligence, in fact, will be the closest we can get to something like 'pure thought', a purely logical thinking. As for writing prose, computers can already do this; cf. novelist and literary editor John Lanchester writing in the London Review of Books (March 5, 2015). In a review of two books on automation and the impact of robots on society, titled "The Robots Are Coming", Lanchester quotes a financial report generated by a computer and concludes that many white collar workers will soon be at risk of being replaced by computers (see quote §2 below). He comments on the report:

  > …the fact to concentrate on is that it wasn't written by a human being. This has been a joke or riff for so long – such and such 'reads like it was written by a computer' – that it's difficult to get one's head around the fact that computer-generated news has become a reality.

- Cognitive science sees the mind-brain as a computer, an 'information-processing' system. In 1995 philosopher Ned Block published an article titled "Mind as the Software of the Brain". The title expresses the prime assumption on which cognitive science is based. Mind emerges from or consists in the processing the brain's circuitry engages in. From this perspective, the idea that science can model or actually reproduce the mind's in its 'processing' operations is something like a necessary item of faith. As Drew McDermott, a professor in the Artificial Intelligence department at Yale, puts it

> … in the current state of cognitive science it must be an open possibility that a computational model of the human mind will be developed that accounts for at least its nonphenomenological properties.
>
> (On The Claim That A Table-Lookup Program Could Pass The Turing Test, Drew McDermott, *Minds and Machines*, 2014)

> This doesn't mean cognitive scientists believe the brain is like an electronic computer or that the software constituting mind takes the same form as the programs running on electronic computers. Cognitive science doesn't actually think the brain has its own programming language or languages – although it does entertain the idea of a fundamental neuronal code – and nor does it hold that the 'circuitry' made up of the network of the nerve cells (neurons) in the brain is anything like electronic circuitry. Instead it assumes the brain works on the model of 'neural networks', also known as 'connectionist systems'.[2] The objections you raise against the possibility of human speech apply strictly to models of the brain based on conventional computing. Neural networks are exempt from objections of this type.

There are a number of arguments one can put forward in answer to these objections. I won't try and rehearse them here, partly because I am running out of time but also because at this point the counter-arguments tend to involve various technical and semi-technical philosophical and computational considerations. In a sense, though, considerations and arguments from philosophy and the theory of computing are no longer relevant or at least carry no weight when set against the near-universal public acceptance of the idea of artificial intelligence.

Two factors in particular make the debate more a matter for social psychology (i.e. the psychology of belief and opinion in both the public sphere and the academic field) than reasoned discussion. First, and perhaps most decisive, is the fact that in recent years leading figures from the fields of science and computing (mainly the computer industry) have issued warnings about the possible dangers of developing an artificial intelligence. The most prominent example is an open letter published in January this year signed by some very distinguished figures in these fields. In an article in the New York Review of Books (see quote §4 below), Sue Halpern described the letter's message as a "call to conscience":

> … a group of more than one hundred Silicon Valley luminaries, led by Tesla's Elon Musk, and scientists, including the theoretical physicist Stephen Hawking, issued a call to conscience for those working on automation's holy grail, artificial intelligence, lest they, in Musk's words, "summon the demon."

The effect of these kinds of pronouncements is in fact more rhetorically persuasive than any positive argument for AI could be. For the letter takes it for granted that building an intelligent machine is not only a feasible technological possibility, but also one that will be shortly realized.

[2] How this radically different conception of computing squares with the computing electronic computers do is something cognitive science rarely addresses. In our view, it represents a significant flaw in cognitive science's fundamental picture of mind and brain (but this is not the place to discuss the problem).

The other factor in making the topic of AI hard to approach in an objective fashion is the widespread representation of robots in science fiction films. Two recent examples are the films *Ex Machina* and *Her* (see attached review by Daniel Mendelsohn, "The Robots are Winning!"). The first portrays a 'humanoid' robot (what used to be called an 'android') in a highly naturalistic fashion and with what appears to be a great deal of scientific plausibility; the second involves a fully intelligent operating system, capable of conducting conversations with people. *Her* is described as a "romantic science fiction comedy-drama" but it takes the idea of an operating system that can speak (using the very expressive voice of Scarlett Johansson) perfectly seriously. All the reviews I've read of these films accept their basic premise – that an intelligent machine capable of thought and speech will be built sooner or later.

So the position of the AI sceptic is an extremely difficult one to sustain: it effectively commits the sceptic to the view that a large number of distinguished scientists and experts in computer technology are mistaken and in the grip of a delusion. Actually, it's not quite as dire as that: there are quite a few people, a handful of them working in the academic field of artificial intelligence, who are not as sanguine about the prospects of building an intelligent machine. But their views receive little or no attention. The belief in the possibility of creating an intelligent machine now has the status of orthodox wisdom and established scientific truth. In the current climate of opinion, expressing doubt about AI comes close to taking a heretical stand.

Robots can't speak

I planned to go to town on this theme, but for the moment I'll have to make do with some brief notes.

There are two senses in which robots can't speak:

1. In technological reality and practice: even the most advanced robots today have no capacity whatsoever for speech, let alone understanding. In so far as they go through the motions of producing speech, this is a kind of simulacrum or impression of language ability which involves the use of pre-determined or 'canned' spoken sentences.

   It is highly significant that AI enthusiasts and believers are generally silent on the issue of robot speech or natural language processing (NLP). Manufacturers of robots like Honda rarely, if ever, mention the absence of any degree of language ability in their machines. Academic writing on AI does sometimes acknowledge that the task of programming a computer to speak and understand is an area no one has made any progress in, but more generally tends to gloss over the problem or dismiss it with the assumption that it will sort itself out as computers become faster and more powerful.

2. Speech and understanding, I argue, presuppose a speaker, a person or some other form of living creature who speaks and understands speech. The idea of a machine speaking is therefore unintelligible unless we endow the machine beforehand with something like personhood or a self; that is, unless we imagine the machine as a living (non-biological) being. The alternative is the incoherent notion of speech

without a speaker. There is a lot more to be said on this theme, but, again, I've run out of time and will have to save it for another document.

A very useful way of getting an idea of what is at stake in the area of computer speech and understanding is to try some of the online 'chatbots'. These are programs that employ various software tricks to give an impression of intelligent conversation. The first such program, ELIZA, was written by a professor at MIT, Joseph Weizenbaum, in 1966. His purpose was in part to demonstrate the ease with which it possible to deceive people that a machine has intelligence. ELIZA consists of a fairly short program (no more than few hundred lines of program 'code') and a database (or just a list) of ready-made phrases and keywords that the program is cued to pick up on.

You can try Eliza online at

www.manifestation.com/neurotoys/eliza.php3

It's also worth trying the rather more advanced chatbot program Mitsuku at

www.mitsuku.com

Mitsuku won the Loebner Prize competition in 2013. It uses a large database of ready-made answers and keywords, and has also been designed to respond to questions that test genuine understanding (see quote §3). In other words, the programmer of Mitsuku has collected the sort of questions people ask to expose a chatbot as a chatbot program (and not a real intelligence) and has built in routines to handle questions of this type. It is quite difficult to catch Mitsuku out and when you do ask a question it is not prepared for it has a number of ready-made phrases to cover up its lack of understanding. As you can see, it's difficult to avoid anthropomorphising the program – I shouldn't really talk of "lack of understanding" because a piece of software is not the sort of thing that could have understanding.

For example, I typed in the question: "What is the opposite of crooked?" Mitsuku answered: "straight".

Then I typed the question: "What is the opposite of obscure?" Clearly the word "obscure" is not in its database, but it is programmed to give a pseudo-intelligent response to sentences with the phrase "what is the opposite of" in them. It answered: "anti-obscure?"

Mitsuku had no problem with the type of question people often ask as a way of testing understanding, "How many legs does a cat have?" and answered "I would guess that one cat has four legs?". But when I asked it, "How many walls does a room have?", it replied, "About a million."

Quotes

1. Nor Shall My Sword, F. R. Leavis, Chatto & Windus, 1972; 'Literarism' and 'Scientism': the Misconception and the Menace, p. 142:

> It's the general blankness in the face of the issues that's so discouraging. I was, I confess, a little amused when, sitting at a formal lunch next to the director of a City Art Gallery, I was told by him, in the tone of one saying something very impressive: 'A computer can write a poem.' I replied, very naturally, that I couldn't accept that, adding that it was one of the things that I *knew* to be impossible. When he responded by being angry, fierce and authoritative, I reflected that he was German, if an émigré, and that in any case his business was *Kunst* and he hadn't said that a computer could paint a work of art. The other occasion on which I was confronted, point-blank, with the preposterous and ominous claim, which by then I had discovered to be pretty current, it made a profound impression on me. The testifier was a philosopher, a lady and cultivated; her place and conditions of residence gave her access to a friendly computer laboratory. She had taken advantage of the opportunity, I gathered, to develop an intense experimental interest: 'It's incredible,' she said, 'what a computer can do; it's awfully fascinating; you know, a computer can write a poem.' I couldn't let that pass; with the appropriate urbanity I said: 'Well, "poem" means different things.' There was no Teutonic anger this time. There was a sudden descent, a heightened nuance of pink, a concessive philosophic laugh, and then; 'O well, yes; but it's great fun.'
>
> That any cultivated person should *want* to believe that a computer can write a poem! – the significance of the episode, it seemed to me, lay there; for the intention had been naive and qualified. It *could* be that because of the confusion between the different forces of the word 'poem'. And yet the difference is an essential one; the computerial force of 'poem' eliminates the essentially human – eliminates human creativity. My philosopher's assertion, that is, taken seriously, is reductive; it denies that a poem is possible – without actually saying, or recognizing, that. If the word 'poem' can be used plausibly in this way – and by 'plausibly' I mean so as to be accepted as doing respectable work – so equally can a good many other of the most important words, the basically human words. Asked how a trained philosophic mind in a cultivated person could lend itself to such irresponsibility, I can only reply that the world we live in, the climate, makes it very possible.

2. The Robots Are Coming
John Lanchester
LRB, Vol 37 No. 5, March 5 2015

> And it's not just manual labour. Consider this report from the Associated Press:
>
> > CUPERTINO, Calif. (AP) Apple Inc. (AAPL) on Tuesday reported fiscal first-quarter net income of $18.02 billion.
> >
> > The Cupertino, California-based company said it had profit of $3.06 per share.
> >
> > The results surpassed Wall Street expectations … The maker of iPhones, iPads and other products posted revenue of $74.6 billion in the period, also exceeding Street forecasts. Analysts expected $67.38 billion.

> For the current quarter ending in March, Apple said it expects revenue in the range of $52 billion to $55 billion. Analysts surveyed by Zacks had expected revenue of $53.65 billion.
>
> Apple shares have declined 1 per cent since the beginning of the year, while the Standard & Poor's 500 index has declined slightly more than 1 per cent. In the final minutes of trading on Tuesday, shares hit $109.14, an increase of 39 per cent in the last 12 months.

We'll be returning to the content of that news story in a moment. For now, the fact to concentrate on is that it wasn't written by a human being. This has been a joke or riff for so long – such and such 'reads like it was written by a computer' – that it's difficult to get one's head around the fact that computer-generated news has become a reality. A company called Automated Insights owns the software which wrote that AP story. Automated Insights specialises in generating automatic reports on company earnings: it takes the raw data and turns them into a news piece. The prose is not Updikean, but it's better than E.L. James, and it gets the job done, since that job is very narrowly defined: to tell readers what Apple's results are. The thing is, though, that quite a few traditionally white-collar jobs are in essence just as mechanical and formulaic as writing a news story about a company earnings report. We are used to the thought that the kind of work done by assembly-line workers in a factory will be automated. We're less used to the thought that the kinds of work done by clerks, or lawyers, or financial analysts, or journalists, or librarians, can be automated.

3. *The Status and Future of the Turing Test*
James H. Moor
Minds and Machines, no. 11, 2001

During the Loebner 2000 contest there was an unofficial eleventh 'judge' who asked some questions and gave a couple of commands to all of the respondents both humans and computers. This 'judge' posed these queries solely to gather information and was not involved in the scoring. The queries were fixed in advance around the three areas: understanding, reasoning, and learning. Here were the questions and commands posed:

Understanding:
1. What is the color of a blue truck?
2. Where is Sue's nose when Sue is in her house?
3. What happens to an ice cube in a hot drink?

Reasoning:
4. Altogether how many feet do four cats have?
5. How is the father of Andy's mother related to Andy?
6. What letter does the letter 'M' look like when turned upside down?

Learning:
7. What comes next after A1, B2, C3?
8. Reverse the digits in 41.
9. PLEASE IMITATE MY TYPING STYLE.

> Understanding, reasoning, and learning (URL) are not, of course, independent categories. If one understands something, most likely one has learned it at some time and probably done some reasoning about it. Learning in turn requires some understanding and so forth. These are intended as common sense categories that are connected and jointly cover a significant region in the domain of ordinary intelligence. As used here, understanding is characterized by a virtually instantaneous grasp of a situation. One doesn't have to think about the issue very long, at least on a conscious level; the analysis of the situation is apparent. Reasoning, requires a few seconds of putting the pieces together, but the assembly need not be difficult. Finally, learning requires various skills such as making an induction, following instructions, and imitating an example.
>
> [None of the chatbot programs in the Loebner 2000 contest could answer the questions above. In subsequent Loebner Prize contests questions of this kind (that test genuine understanding) were prohibited.]

4. How Robots & Algorithms Are Taking Over
Sue Halpern
NYRB April 2, 2015
Review of "The Glass Cage: Automation and Us" by Nicholas Carr

> But since automation also produces quicker drug development, safer highways, more accurate medical diagnoses, cheaper material goods, and greater energy efficiency, to name just a few of its obvious benefits, there have been few cautionary voices like Nicholas Carr's urging us to take stock, especially, of the effects of automation on our very humanness—what makes us who we are as individuals—and on our humanity—what makes us who we are in aggregate. Yet shortly after *The Glass Cage* was published, a group of more than one hundred Silicon Valley luminaries, led by Tesla's Elon Musk, and scientists, including the theoretical physicist Stephen Hawking, issued a call to conscience for those working on automation's holy grail, artificial intelligence, lest they, in Musk's words, "summon the demon." (In Hawking's estimation, AI could spell the end of the human race as machines evolve faster than people and overtake us.) Their letter is worth quoting at length, because it demonstrates both the hubris of those who are programming our future and the possibility that without some kind of oversight, the golem, not God, might emerge from their machines …

Simon Beesley
simonbeesley@clara.co.uk
July 2015