

Philosopher's Zone

on ABC Radio National

Minds and Computers

12 May 2007

Could a computer have a mind? Is our mind a sort of computer? And what do we mean by 'mind' anyway? This week, we talk about these mind-numbing issues with Matt Carter, Fellow of the Philosophy Department at Melbourne University.

[Hide Transcript](#)

Transcript

Alan Saunders: Hi, I'm Alan Saunders and this is *The Philosopher's Zone*.

Now tell me, are you a computer? And could there ever be a computer that was in some way like you? Well, *Minds and Computers* is an introduction to the philosophy of artificial intelligence by Matt Carter, who is a Fellow of the Philosophy Department at Melbourne University. And when I spoke to him, we began with man-made life from fiction: Frankenstein's creature, for example, and the robots invented by the Czech playwright, Karel Capek. How do they differ from artificial intelligence?

Matt Carter: Well I suppose the main difference is that in earlier works of fiction these creatures were generally cashed out in terms of something like mechanical men, it was the metaphor of the day, whereas in the latter half of the 20th century we see a marked move to something like a computational artificial intelligence, a computer that has a mind in the sense that we take ourselves to have minds.

Alan Saunders: Why is this an interesting question? Is it an interesting question because we think that perhaps we could develop computers that are like us in some intellectual respect and to whose rights we will perhaps have to give recognition? Or is it because we think that the computational model will tell us something about our own minds?

Matt Carter: It's an excellent question, and I think the answer is both. There's a sense in which we really hope to understand our own minds better through this kind of computational understanding, and certainly the computational theory of mind is currently by far the most dominant theory in the philosophy of mind and the culture of sciences broadly. But there are also a number of people working on strong artificial intelligence projects, and the ultimate goal of those projects is to produce man-made artifacts that have minds in precisely the same sense, or some very similar sense, in which we take ourselves to have minds.

Alan Saunders: Well let's look, as you do in the book, at various images of the mind, various views

of the mind that have developed over the centuries, and the most obvious one to begin with, is the dualism most famously postulated by the 17th century French philosopher René Descartes. Now Descartes' view is - and it's slightly more subtle than it's often made out to be, but essentially Descartes' view is that there's a mind and there is matter, and the two are separate. And I find it fairly easy to see the appeal of that for a variety of reasons. One is that there are aspects of my mind which don't seem to me to have much to do with my body; the other of which is that, if my mind is independent of my body, then perhaps my mind will go on when my body dies. So what do you see as the appeal of, and the arguments against, this Cartesian dualism, as it's called?

Matt Carter: OK, there are a number of arguments against Cartesian dualism, and it's certainly not a new theory in the history of ideas. It's where I begin telling the story of the history of philosophy of mind, but it's an idea that goes back to Plato, but Descartes give us a very rigorous formulation of it. It's a view about metaphysics primarily, a view that the universe is composed of two different kinds of stuff, not just different in the sense that tables and chairs are different to things like heat and light and sound, but more markedly different in the sense that there's the physical universe, the kind of thing that we expect physics to tell us about, and this includes things like heat and light and sound; and then there's the non-physical universe. It's completely, fundamentally, substantially distinct from the physical universe, and the Cartesian view was that minds inhabit this latter realm of the non-physical, but minds, if you will, in a kind of one-to-one correlation with bodies which inhabit the physical realm. So the most obvious and potent objection against this Cartesian view, Cartesian dualism, or interactionist dualism, as it's sometimes called, is what's called 'the problem of interaction'.

Now one wants to maintain, as a Cartesian dualist, that the mind has a causal effect on the body and the body has a causal effect on the mind. I mean this is commonsense. I can think about raising my arm and my body moves in such a way that my arm raises, and I can perhaps bang my hand and this causes a mental event of pain. So it's fairly intuitively and in a commonsense fashion, clear that there is this relation between minds and bodies. Any theory has to account for this. The problem with the Cartesian view is that it's not at all clear how some non-physical substance or some non-physical entity can be causally efficacious in the physical realm, and indeed this contravenes the basic laws of thermodynamics, and this is not something that one wants to give up lightly.

Alan Saunders: OK, well, if we're not going to have that sort of dualism, another alternative is we might say that psychology, the science of the mind, it's just about observing behaviour, and this is what's called behaviourism. Now you distinguish between psychological behaviourism, which is a method of studying people, and you mention in that context an experiment performed by I think it was B.F. Skinner, on a boy, an 11 month old boy, who was given a fear of rats. He would look at the rat and stroke the rat out of curiosity, and then Skinner would clap his hands or make a noise and as a result of that rather threatening sound, the boy developed a fear of rats, and of various other things, and when I read that I thought Where's an Ethics Committee when you need one?

Matt Carter: Precisely.

Alan Saunders: But that's psychological behaviour, isn't it, it's a way of studying the human mind, but there's also philosophical behaviourism; tell us about that.

Matt Carter: You're absolutely right to distinguish between the two. Philosophical behaviourism, is a substantive theory of what mental states are. So it's a philosophical theory of mind and it can be seen really as a reaction to dualism and historically situated with the growth of psychology as an independent discipline from philosophy, which is something that's only happened in the last 100 years or so. Behaviourism can be seen as an attempt to formulate a philosophical theory of mind that provides the kind of rigorous philosophical foundation that we want from a philosophy of mind, but also provides direction for this fledgling science, this science of the mind. So the thought was that, rather than conceptualising mental states as something going on in some spooky immaterial spirit/mind kind of stuff in a Cartesian fashion, we should just understand mental states as behaviours naively, but really we need to understand these things as dispositions to behave. So the behaviourist analysis of mental states says that when I'm in a mental state, what that actually means, what we mean when we say of someone that they're in a mental state, is precisely that: they're disposed to behave in particular ways.

Alan Saunders: That seems to me to be very obviously wrong. I mean, it seems to me that there is an interior quality to my mental state. Possibly if I'm happy, you can tell that I'm happy because I'm bouncing around. If I'm unhappy, I'm just sort of sitting there. I might be looking a bit glum, but I might not be looking like anything. But I'm unhappy because that's the quality of my interior state, and behaviourism misses that completely.

Matt Carter: No, I think you're exactly right. That's the major problem really with behaviourism. It really misses out on what's fundamentally important about having the kinds of minds that we have, and it misses the fact entirely that there is something that it's like to be in any particular mental state. This is something that just can't possibly feature in a behaviourist analysis because it's not something that can be loaded in to the sort of dispositional paraphrases that the behaviourist analysis aims to give of mental state terms.

Alan Saunders: OK, well another possibility, and let's be patriotic here, because it's an Australian idea, is that we simply identify states of mind with states of the brain. This is what's still called Australian materialism. Tell us about that?

Matt Carter: This is a theory that initially held out a lot of hope. So we're talking the '50s and up until about the early '60s or so, and the main names associated with this theory are J.J.C. Smart, 'Jack' Smart, and U.T. Place, Alan Place. The idea here was that we want to maintain the sense in which behaviourism was leading us in the right direction for a philosophical theory of mind. We want our philosophical theory of mind to still give methodological direction for the empirical disciplines, but we want to cash it out a little bit better, a little bit more sophisticatedly and we want to take into account what's just starting to happen in a more sophisticated fashion about this time, namely various kinds of neural imaging. We're starting at this point to be able to use various

methodologies to see what's going on in brains. The force at heart of Australian materialism is that mental states, types of mental states rather, just are types of neural states, so when for instance I'm thinking about chocolate ice-cream, well that just means that I'm in a particular type of neural state, and initially, this is a very promising theory.

Alan Saunders: I suppose again, I want to say that a fact of neuro-physiology and a fact about my mind, however closely connected they are, are not the same thing. And I know at least one philosopher, the Austrian-American philosopher, Paul K. Feyerabend, says, Well look, it's like lightning. The stroke of lightning is one thing, and the actual physics of lightning is another, but that's all there is to it, the physical story is all there is.

Matt Carter: And I think an Australian materialist would certainly agree with you there. But once again, we're missing, if you will, this important element of cognition, namely that it's something, there is something that it's like to be in these mental states. So while it's tempting to take this kind of analysis, it doesn't seem to flesh out the picture fully, in the sense that we require from a philosophical theory.

Alan Saunders: So let's move on to the age of computing. How do computers and how might artificial intelligence help us to answer questions about the mind?

Matt Carter: There's a sense in which it provides us with not just a metaphor, but an analysis that can be understood through metaphor of what minds are, which really helps us understand how we might go about creating artefacts that have minds in this sense, but also help us understand the properties of our own minds, and they help us in particular to understand the relation between brains and minds. So this new theory, this currently dominant theory, which is perhaps not so new now, but new in terms of history of ideas, sees mentality, sees cognition, as analogous to software, and understanding it in these terms, we can get a really good handle on the relation between brains and minds. Brains, on this view, they're cognitive, computational hardware, biological computational hardware, the term of art in the discipline is called wetware, so we think of brains as wetware and we think of minds as the software that runs on this computational hardware, and one of the really nice advantages of this view is it allows for lots of different kinds of things to have minds.

Alan Saunders: Let's look at what we're talking about when we're talking about computers, because they've been around for a long while, and the original theoritisation of computers, I mean you had people who were trying to create calculating machines a long time ago, but essentially they were doing spaghetti with wires. But then you have Alonzo Church and Alan Turing who are the great theorists, early theorists of the computer. So what are they giving us and how is it useful?

Matt Carter: Well, what they gave us in essence is something we take for granted very much these days, the computer as we commonly understand it, the thing that sits on most people's office desks. They gave us the computational theory that allowed for the possibility of the generation of these devices, and they gave us an understanding of computation as syntactically structured symbol

manipulation. So computation essentially, the very short and hand-waving answer, is simply taking a string of syntactically structured symbols and performing some content-respecting operation on that to give some other syntactically structured string of symbols. And one can do extraordinarily powerful things with very simple methods.

Alan Saunders: OK, now you've used the word 'syntax', which is a linguistic term which suggests that the nature of language is very important to this discussion. So how does the way in which we use and interpret language enter into the discussion?

Matt Carter: That's an excellent question. The information from linguistics on natural language processing really helps very much in informing philosophical theories of mind. Now Chomsky reconceptualised syntax, or grammar, if you will, for present purposes, as a generative process. He thought, well, there are a number of rules, and basic states that allow us to generate all the possible grammatically correct strings of a language, and thinking about grammar in these terms makes it very clearly a computational task, and it helps us give an understanding of natural language processing that's computationally implementable.

Alan Saunders: Now there are a number of objections to the notion of the brain as a computer, and possibly the most famous one comes from the American philosopher John Searle, and it's the Chinese room argument. Tell us about that.

Matt Carter: This is a fascinating thought experiment. Thought experiment is very important in the philosophy of mind. Thought experiments test logical claims. Now by their very nature thought experiments very often describe outlandish situations, and it's very tempting on first presentation of a thought experiment to think, well that just couldn't at all be possible, it's not physically realizable, so why should I entertain such an idea. The idea of thought experiment is to consider, Well what would follow if the described situation were the case. Now the Chinese room thought experiment is very well known in philosophy of mind, and it's an argument, and to put it in slogan form, that syntax is not sufficient for semantics. So syntactic manipulation is the moving around of symbols, but syntax of a language we can think of for present purposes, as just like the grammar of a language. The thought that the Chinese room thought experiment primes is that no amount of doing this is going to tell you what these things mean. So if natural language processing is just syntactical manipulation, if it's just moving around our strings in order to meet conditions of grammaticality, then there's no way in which the meaning of these terms could ever come out of the system. And the thought experiment goes like this:

Imagine that someone asks me to sit in this very large room to help them out with an experiment that they're working on. Now I walk into the room, and I see that it's enormous, and it's covered with shelves, and these shelves are covered with books and all of the books are numbered. And on a table in front of me, the sole table in the room, and a sole chair, I see a book that's open. And the researcher tells me that this is what's going to happen. Someone is going to pass a string of symbols under the door, and I'm going to have to find whereabouts in the book this string of symbols appears.

Now the page on which it appears, right next to it there'll be another string of symbols and a number, and all I have to do is write out the second string of symbols, pass it back through the door and then the number tells me which book I should choose from the shelves next. And I do this all day. And I have no idea what I'm doing, really, I'm just copying out symbols in response to receiving other symbols.

Now at the end of the day, I'm told that outside the door was a Chinese native speaker and what they've been passing me have been symbols in Chinese, and what I've been doing is passing back out reasonable responses to these questions that have been passed under the door. Now I was able to do this by virtue of the fact that all the possible conversations I could have in Chinese were encoded, so all I was doing was moving through conversational states by performing syntactically structured operations on symbols. But I don't thereby have any understanding of Chinese, and so no amount of this kind of syntactic symbol manipulation is ever going to tell me about the meaning of the terms, or if you will, syntax is insufficient for semantics.

Alan Saunders: Well the American linguist, Stephen Pinker, would say, Well in fact that is what happen in the human mind, it just happens so quickly we don't notice that we're looking things up in a book, but that is actually how the mind works.

Matt Carter: I think one can sensibly maintain that kind of view, but the important distinction between what I'm doing inside the Chinese room and what's happening inside my head when I move about in the world, is precisely that thing that I just mentioned: I'm moving around in the world, I'm embodied, I have sensory apparatus that connects me to the external world and it's by virtue of these relations to the external world, this embodied sensory experience, that the terms that I use come to have meaning. So this is something that I don't have in the Chinese room. If the Chinese speaker asked me for instance, Would you like a hamburger? They put this in the requisite symbols, and I passed out symbols that said, No thank you I'm not hungry, this is completely divorced in the first instance from my beliefs and desires about things like hamburgers and hunger, but in the second instance and quite importantly, divorced from any relation to actual hamburgers. Whereas when I'm moving around in the world and this kind of question comes up, hamburgers are salient in the environment, if you will, or at least I've run across them in the past, and it's by virtue of this kind of relation to objects in the external world that my terms come to be conferred with me.

Alan Saunders: Well this relates to a question I asked you at the beginning of this conversation, which was, not what artificial intelligence tells us about our own mind, but what these investigations tell us about artificial intelligence, and it is in the nature of a computer, of an artificial intelligence that it isn't embodied. So, however clever it is, and even if it develops a personality like Hal in 2001, it's not embodied, and that makes a huge difference, doesn't it?

Matt Carter: I think you're exactly right. It does indeed make a huge difference, and if we really want to achieve strong artificial intelligence, I think embodiment is a necessary condition on the possibility of such development. I think this is the intuition that one derives from the Chinese room

thought experiment. Most importantly, that if one wants to develop a mind in the sense that we take ourselves to have minds, these things must be embodied.

Alan Saunders: This is highly speculative, but does it have to be an embodiment similar to ours? Does it have to be, as it were, soft and squashy?

Matt Carter: Mmmm...it's a very good question. I don't think that falls out of the Chinese room experiment. I think what falls out is that one needs some kind of mediated relations with the external world, and that mediation might occur in a number of different ways, but you might think that in order to have qualitative experiences that are like ours one needs human-type sensory apparatus. One might think that we have the kinds of minds we have by virtue of being embodied in the way we are.

Alan Saunders: The acknowledgments to your book end with the sentence, 'Thank you to Mia and Linus for being adorable, and a million thank yous to Sue for being wonderful.' Who are Mia and Linus, by the way?

Matt Carter: Mia and Linus are our pets;

Alan Saunders: I thought they might be.

Matt Carter: Mia is my Staffordshire Terrier and Linus is our cat. And I believe that they both have minds, clearly not in the sense that we have minds, and they clearly don't have the full range of cognitive apparatus that we have, but nonetheless I think it's implausible in the extreme, and I think any pet owner would agree with me, implausible in the extreme to deny that they have a simple belief desire psychology.

Alan Saunders: Actually, that's a view to which, having had a close relationship with a cat over the last couple of years for the first time ever in my life, I have come to be persuaded of. Buzz, the cat in question, has hugely educated me I think in how minds work. But, thank you to Mia and Linus for being adorable, and a million thank yous to Sue for being wonderful. These are the words of an embodied mind, aren't they?

Matt Carter: Indeed, precisely.

Alan Saunders: And so given that you're an embodied mind, how do you end the argument? What is your summing up?

Matt Carter: My summing up is that I don't see any in-principle barrier on the development of strong artificial intelligence, but there are some rather severe technological barriers, and there are a number of open questions, and it's not at all clear that we would be able to develop an artefact, a man-made device that would have minds in the sense that we take ourselves to have minds. If we'd want such a thing, then it's still by far and away the easiest method to develop is to simply have a child.

Alan Saunders: Have you got a child yet?

Matt Carter: No.

Alan Saunders: I'm sure that will be an education too.

Matt Carter: No doubt.

Alan Saunders: And have you changed the view of your own mind and of how your mind operates?

Matt Carter: I think I've learnt a lot of useful things about my mind throughout my studies and I think a number of these things were surprising, in particular learning about linguistics and natural language processing, and the extraordinary number and complexity of the computational tasks that need to be achieved to do what you and I and all our listeners are doing right now, namely process natural language. It's really quite surprising and enlightening and tells us a lot about the structure and nature of the kind of computational hardware that we must have.

Alan Saunders: Matt Carter, it's been a pleasure, thank you very much for joining us.

Matt Carter: The pleasure's been mine, thank you, Alan.

Alan Saunders: And Matt Carter's book *Minds and Computers: An Introduction to the Philosophy of Artificial Intelligence*, is published by Edinburgh University Press.

The show is produced by Polly Rickard, with technical production this week by Ann Marie Debettencourt. I'm Alan Saunders. And I'll be back next week with another *Philosopher's Zone*.

Guests

Matt Carter

Fellow, School of Philosophy

University of Melbourne

<http://braininavat.net/weblog/index.htm>

Publications

Title: *Minds and Computers: An Introduction to the Philosophy of Artificial Intelligence*

Author: Matt Carter

Publisher: Edinburgh University Press

URL: http://eup.ed.ac.uk/edition_details.aspx?id=12278

ISBN 978 0 7486 2099 9 (0 7486 2099 0)

Presenter

Alan Saunders

Producer

Polly Rickard



Saturday 1.35pm
repeated Monday 1.35pm

Presented by
Alan Saunders

 Philosopher's Zone | Radio National | Programs A-Z