

Conversation With Peter Tse: Artificial Intelligence

[MUSIC PLAYING]

MARCELO GLEISER: There is a lot of debate now about artificial intelligence, and if we humans are going to be able, at some point, in building machines that are not just following rules in a program, like the ones that we have now, but actually have some level of self-awareness. What's your position on that? Do you think that this is something that is achievable? Do you think it's a pipe dream?

PETER TSE: Well, I think, given computers as they are now, it will never be accomplished. That doesn't mean that in future versions of computing systems that it can't be accomplished. I mean, so the brain is just manifestly not a computer. So the dominant metaphor of my field-- neuroscience, cognitive neuroscience, psychology-- is that the brain is a kind of computer. And this just fails. Right? Because there is no software/hardware distinction in the brain. No computer is rewiring itself on a millisecond timescale. And computers are not conscious. So we know that our governing metaphor is false.

And I would argue that computers as they are currently realized are very algorithmic. And algorithms-- one simple way to think of it is as a thread of decisions. There's this input. And then there's yes or no decision, and then a single output, whereas the neurons are just radically different from that. They're not just taking a single thread of input. They're taking 10,000 inputs, integrating them, and then setting hundreds or thousands of inputs out.

So I'm skeptical. And yet, we are physical systems. Evolution discovered the path to consciousness. The fact that we're experiencing now is proof that it's possible for a physical system to be conscious. But I look at the latest greatest stuff in AI, and I'm very skeptical.

I mean, so the latest greatest thing right now would be these deep learning networks. People first were happy about 10 layers. And now there's hundreds of layers. Right? And these things are impressive. They can, for example, not only, say, recognize that this painting is Manet or a Monet or van Gogh. They can actually take a photograph and rework it to look like a van Gogh. So that's impressive. But that's nowhere near what we are doing.

So what we are doing is we have mental models. We're not just taking input and processing it in a bottom-up way. We have mental models of how things work. And we can use those mental models to operate not only in the world, like driving a car, I can use those mental models in my internal virtual reality, which is also conscious. So there's consciousness of the outside world. Then there's my internal virtual reality, which can also be conscious. You might call it daydreaming, fantasy, dreaming, working memory. And I can use these mental models to carry out operations, like play things out. I see nothing like that in any computer system now. There's nothing like mental models.

That doesn't mean that they can't eventually be instantiated. But my guess is by the time that computer science realizes anything like that, computers themselves will look radically different. They'll end up looking a lot more like the brain than this von Neumann architecture.

MARCELO GLEISER: OK, and so to finalize, what do you think of this whole debate about assuming some sort of true artificial intelligence is possible in the future? Should we be afraid?

PETER TSE: OK, so maybe some people have seen this movie, Ex Machina, or there's this other one. I mean, I know that there's a lot of worry about robotization, automatization. And in terms of economics, I mean, that's something to worry about. Then the question is, let's say that we create robots that can actually do things. Let's say that we want to create a company that will make robots that are essentially slaves. So in order for that to be possible, I have to give a command to this robot, like, do the dishes. Yes, and they'll do it. And so a general purpose robot like that, I think, would only be able to really do that if it had mental models that we have. Right? Because there's so much ambiguity. If I say change the tires, well, you have to know change the tires on the car, and which car, the car with the broken--

Now, where do our mental models come from? Partly, they come from our consciousness. So the fact that I know what it's like to feel pain means I understand what it's like for you to feel pain. Because I have my mental model of my pain, I can generalize that. So the puzzling, or the funny, or ironic thing is that if we get to the point where we create robots who can actually do useful things and become our slaves, they would probably have to also be conscious in order for them to understand what we're talking about. But if they're conscious, then we face all the normal problems of human slavery. It would be immoral because they can feel pain. And so in the end, it would be just as immoral to have a conscious robot that can feel pain that we're enslaving as to enslave a human being.

[MUSIC PLAYING]

Conversa com Peter Tse: Inteligência artificial

[MÚSICA TOCANDO]

MARCELO GLEISER:

Hoje em dia há muito debate sobre a inteligência artificial e se os seres humanos serão capazes de, em algum momento, construir máquinas que não sejam apenas seguidores de regras em um programa, como as que temos hoje em dia, mas que tenham algum nível de autoconsciência. Qual a sua posição sobre isso? Você acha que isso é algo viável? Você acha que é um sonho inalcançável?

PETER TSE: Bem, eu acho que com os computadores que temos agora, isso nunca será possível. Isso não significa que em futuras versões de sistemas de computação isso não possa vir a se tornar real. Quer dizer, o cérebro não é simplesmente um computador. A metáfora dominante da minha área, a neurociência, a neurociência cognitiva, a psicologia, é que o cérebro é um tipo de computador. E isso simplesmente falha. Não é mesmo? Porque não há distinção de software/hardware no cérebro. Nenhum computador se reconecta em uma escala de tempo de

milissegundos. E os computadores não são conscientes. Por isso, sabemos que nossa metáfora governante é falsa.

E eu diria que os computadores como são atualmente idealizados são muito algorítmicos. E os algoritmos, uma maneira simples de pensar nisso, é como um fio de decisões. Há uma entrada. E, então, há uma decisão sim ou não e, em seguida, uma única saída, enquanto os neurônios são radicalmente diferentes disso. Eles não pegam apenas um único fio de entrada. Eles pegam 10.000 entradas, integrando as e, em seguida, definindo centenas ou milhares de outras entradas.

Então eu sou cético sobre isso. No entanto, somos sistemas físicos. A evolução descobriu o caminho para a consciência. O fato que estamos experimentando agora é a prova de que é possível que um sistema físico seja consciente. Mas eu olho para as últimas grandes descobertas sobre IA e me encontro muito cético a respeito. Quero dizer,

a descoberta mais recente seria estas redes de aprendizagem profunda. Os primeiros cientistas ficaram felizes com 10 camadas. E agora há centenas de camadas. Certo? E essas coisas são impressionantes. Elas podem, por exemplo, não apenas reconhecer que esta pintura é um Manet, um Monet ou um Van Gogh. Como podem tirar uma fotografia e reformulá-la para parecer um Van Gogh. Isso é impressionante. Mas isso não está nem perto do que somos capazes de fazer.

A realidade é que temos modelos mentais. Não apenas pegamos uma entrada e a processamos de forma crescente. Temos modelos mentais de como as coisas funcionam. E podemos usar esses modelos mentais para operar não só em coisas corriqueiras, como dirigir um carro, podemos usar esses modelos mentais na nossa realidade virtual interna, que também é consciente. Portanto, há uma consciência do mundo exterior. Então há a minha realidade virtual interna, que também pode ser consciente. Podemos chamá-la de sonhar acordado, fantasia, sonho ou memória de trabalho. E podemos usar esses modelos mentais para realizar operações, como jogar alguma coisa fora. Eu não vejo nada assim em qualquer sistema de computador da atualidade. Não há nada como modelos mentais. Isso não

significa que eles não possam ser, eventualmente, instanciados. Mas meu palpite é que se a ciência da computação fizer algo assim, os computadores parecerão radicalmente diferentes. Eles se parecerão muito mais com o cérebro do que com esta arquitetura de von Neumann.

MARCELO GLEISER: Entendi, então só para finalizar, o que você acha de todo esse debate sobre assumir que há a possibilidade de termos algum tipo de inteligência artificial real no futuro? Devemos ter medo? PETER

TSE: Bem, talvez algumas pessoas que viram o filme *Ex Machina: Instinto Artificial*, ou algum outro. Quero dizer, eu sei que há muita preocupação sobre a robotização, a automatização. E em termos de economia, isso é algo para se preocupar. Então a pergunta é, digamos que sejamos capazes de criar robôs que podem realmente desempenhar algumas atividades. Digamos que queremos criar uma empresa que fará robôs que são essencialmente escravos. Para que isso seja possível, eu tenho que dar um comando para este robô, Como, lave os pratos! Daí eles obedecem e fazem isso. Então, um robô de propósito geral como esse, na minha opinião, só seria realmente capaz de fazer isso se tivesse os modelos mentais que temos. Correto? Porque há tanta

ambiguidade. Se eu disser: troque os pneus!, ele terá que saber trocar os pneus do carro, qual carro, o carro com algo quebrado... Agora, de

onde nossos modelos mentais vêm? Em parte, eles vêm de nossa consciência. Então, o fato de eu saber como é sentir dor significa que eu entendo o que é para você sentir dor. Porque eu tenho o meu modelo mental da minha dor, sou capaz de generalizar isso. Então, o intrigante, engraçado ou irônico é que se chegarmos ao ponto de criarmos robôs que podem realmente fazer coisas úteis e tornarem-se nossos escravos, eles provavelmente terão que também estar conscientes para compreenderem do que estamos falando. Mas se eles estão conscientes, então enfrentamos todos os problemas normais da escravidão humana. Seria imoral, porque eles podem sentir dor. E assim, no final, seria tão imoral ter um robô consciente que pode sentir dor do que quando escravizamos um ser humano.

[MÚSICA TOCANDO]