Consciousness Conundrum: Integrated Information Theory vs. Cognitive Consciousness

https://mindmatters.ai/podcast/ep314

Pat Flynn:

Okay, everybody, welcome back to the podcast. My name is Pat Flynn, host of Philosophy for the People teaming up with Mind Matters to continue our conversation with Dr. Selmer Bringsjord on all things AI, and specifically on something called Integrated Information Theory and whether that can explain consciousness. In part one, we just had a sort of general discussion about what our take on AI predicts, the limits of AI, what we should expect in the future. I thought that was really fascinating. I will obviously recommend that people go back and listen to that. But now we're going to take a look at another recent contribution from Dr. Selmer Bringsjord and the title of which is Can Consciousness Be Explained by Integrated Information Theory or the Theory of Cognitive Consciousness? So we're going to take our time with this. This is not an area that I have specialized in. It's a fascinating paper.

So to start out, Dr. Bringsjord, let's do this. This is the way I always like to kind of try to present material that can be quite technical to people who are interested and willing to work hard at understanding certain concepts, but maybe not specialists. Explain this to us like we're five years old, if you don't mind. What is integrated information theory and what is your general argument about in this paper? Let's start really, really broad and then we can dive into the tall grass.

Selmer Bringsjord:

Yeah, I'll try. Five's a little tough.

Pat Flynn:

How about nine?

Selmer Bringsjord:

Right. I can always particularize these sort of lower ages and a challenge of presentation elucidation and teaching now courtesy of my granddaughters, who I, of course appreciate as totally brilliant, but I'm not sure the older one is... I appreciate you moving up to nine because I am not sure at six she is going to get it. Although, I don't know. The stuff on the raw stuff as I think we got into that line of argument and the expansion of it, she is pretty darn good at arithmetic, if you will, number theory, and that's quite amazing.

Pat Flynn:

Yeah. They'll surprise you, I have soon to be six children of my own. I'm constantly surprised at what they can grasp at a young age.

Selmer Bringsjord:

Cool.

Pat Flynn:

But we'll give you nine maybe or up to high school. It'll take whatever you need. Let's just try and get the simple statement out there.

Selmer Bringsjord:

Yeah, the nine is great because to rather perhaps barbarically encapsulate an example we have for the exposition of IIT and in its connection to consciousness and the paper in question is consideration of a robot, actually two robots. And nine year olds probably starting to play at least in virtual environments with robots. It's possible there's science in school early on, which in this age almost inevitably means the science teacher is talking about AI and robots. It's motivating, it's cool, et cetera, the kids love it. So if you look at the challenge of getting yourself or Robby the Robot for the household, because maybe there's some activity in your home. Of course you've got a large family, you can divide all the chores, maybe it all gets done. But there's a lot of drudgery perhaps so you like your robot to come in there and do things in the home.

And that's not just what we primarily see today vacuuming the rugs. It's, "All right, we're all too tired to cook tonight and is daddy going to go out and get takeout?" Which is four miles away and it's cold or whatever, and maybe we just have the robot whip up some food. So cooking, laundry, all kinds of cleaning beyond vacuuming, so dusting. And don't break the cherished family pictures on these tables. And if you find something shiny on the floor, it might be a ring or another piece of jewelry, save that, put it aside, but still get the floor done, et cetera. I think IIT says the following about a comparison between two kinds of household robots that you could get. In one case, Robot Inc. could ship you a robot that has a whole bunch of modules that are separately engineered for the tasks I mentioned. So if it really is going to make a dinner and then do the laundry, it engages two separate modules.

Now, there are probably dedicated algorithms that you would always want to have in the case of a household robot for these two spheres of activity. But these modules that I'm imagining are really separate. There was the laundry team at Robot Inc. and then there was the cooking team and they just decided, "Hey, for all kinds of reasons, we really want to modularize this robot." And then there's the other, maybe Robot Inc. sells two kinds or maybe it's Robot II Inc. creates a robot that is based on a belief that its interoperation across these different activities would be much better because now they can ship a robot that has many less specialists demanded for these dedicated algorithms. If someone says, "Well, I want not a household robot for these purposes, but I need a new robot in my factory floor, what can you do?"

And maybe they Robot II Inc. says, "Well, yeah, easily this more general purpose robot can be exactly what the doctor ordered for you." So in the case of the first robot, it would have lower consciousness, all things being equal than the second robot by virtue of the fact that information in it is integrated across its "intelligence" or its "processing." And this is a humble example, but I think that this diagnosis of severe differential between these two robots having spent some time with the originator of IIT and Christophe Koch as well in a two summer program on technology and consciousness, which we may get a chance to talk about. It was quite an experience. All right, is it a little presumptuous of me? Maybe. But I would say stripping the math out of it, getting down to brass tacks and saying, "Look, what's your basic intuition?"

Not only would they say, but now they would have to say that the robot that's "integrating information" across these modules has a higher degree of consciousness. Most, I don't want to say sane, but I think most thoroughly rational people without a horse in the race hearing what I just said would say, "Are you joking? Neither one of these robots has anything like consciousness? What are you talking about?"

Pat Flynn:

Yeah, that's my point.

Selmer Bringsjord:

Right, exactly. But IIT says something very different. And if you keep scaling this up and look for more movements of information at higher and higher intensity levels and then at a higher and level of integration across the system, again, we don't need to, but we don't want to either get into the mathematics. You are going to have more consciousness. Naveen and I say, "Look, show us first in the robots how if Boston Dynamics sends the first kind of robot to the house. And Google shed Boston Dynamics, so now they have their own robotics division, they send the second robot. But really can you explain how one is more conscious than the other? They're both just doing fundamentally nothing that warrants subscriptions of consciousness." But that's what they would say and that's what they're forced to say. And I say, "Well, it's nonsense." But there's the nine-year-old version of that.

And then you end up, if it's a nine-year-old, you could spook them out. Maybe we're getting close to Halloween so it's appropriate. Maybe I should try. You could say, "Well what if it's not a robot, but if it's your iPhone versus your Samsung smartphone? They are both conscious, you see, but they vary in terms of the readings that you would get if you could take it with IIT." And now since actually Searle himself pointed this out, well then given that ultimately they're made of the same physical stuff going far enough down, do we say that everything that's determinate as a system, maybe we'll put that minimal constraint on it is conscious at a certain... And they have swallowed that.

Pat Flynn:

I would think that they would have to swallow that. Okay, so this is really interesting. And to me it does seem based on a sort of fundamental redefinition of consciousness at the end of the day. And it also seems like it commits them to this idea that conscious is something that is graded, which seems to me not entirely correct. We even think about intentional states. While it makes sense to say that I might have states that represent less, there's no state that is itself less of a representation of anything, right? Either there's a representation or there's not. It seems to me that this is an on-off thing. Same thing with conscious. There might be different feelings of what it is, maybe more feelings than not, but there's either a, what it is likeness or there isn't.

And it seems like there's a certain idea, and help me clarify this. Like I said, I haven't spent a whole lot of time in this, but that we can just take a certain number of... It's entirely quantitative but not qualitative. And I know in your paper you do differentiate between two notions of consciousness. So maybe that's where we should spend some time-

Selmer Bringsjord:

Yeah, no. Important.

Pat Flynn:

... phenomenal consciousness and access consciousness. And I think what's of interest to most people is that first one, right? What it is likeness understanding of consciousness? So yeah, talk to us about that and maybe that'll help clarify the discussion a little bit.

Selmer Bringsjord:

Yeah. So we can back up a little bit. I think pedagogically speaking, for our nine-year-old or twelve-year-old or whatever, we have to, and for the audience-

Pat Flynn:

Might be getting to high school pretty quick here. Yeah.

Selmer Bringsjord:

Oh, we are. And your audience is brilliant. But we still need to back up because we didn't say what we mean by consciousness. And you're right. What they mean by consciousness is so-called P consciousness to use the abbreviation philosophers have really adopted. So it's phenomenal consciousness. And the phrase that you used is right on target, there's something it's like to... And then we would have get pinched severely, get your finger-

Pat Flynn:

Taste coffee.

Selmer Bringsjord:

Right, taste coffee. For me, maybe I should reduce it a bit, but carve a ski turn at high rate of speed with perfect equipment designed for that, race skis. That is just one kind of consciousness, but it's the one that they're definitely after. And I do think with one caveat that you're right, it is there or not, it's either on or not. And to use the phrase I think Stephen Harnaud used many years ago, "There's either someone home or not." The caveat is though, really important. And that is there's self phenomenal consciousness. So you are phenomenally conscious or p-conscious, but you know that you are. You know that you're a subject which is in pain when you are, et cetera. I'm not entirely sure that as we go down through non-human animals, less and less it'll correlate.

So not to identify, but it'll correlate with neuro-computational complexity. As we go down, I'm not sure that it's self phenomenal consciousness. So Harnaud's phrase may be a little bit infelicitous, right? But there's clearly pain being experienced by the creature, okay, whether it has self... And we do this. Unfortunately for us, it makes for sometimes when things go off the rails, mental illness. Because the thing is we have it at two levels all the time. So we say things like to ourselves or we rue the fact that, "Well, I am in pain. Now. I don't want to be in pain in the next hour, so I am going to take acetaminophen."

We reflect on the fact we go one level up and say, "I believe that I'm perceiving my pain." That's two levels. So for them, they've yet to show how any kind of this gradation would work. In cognitive consciousness, our approach to consciousness allows exactly that. And the other challenge to them, I issued this on the spot because the only sort of behavioral prompt we had for coming up with the concept of cognitive consciousness to pit against this and our own measurement scheme, LaMDA. Because they have a measurement scheme, it's called phi. Phi in theory, can measure the level of phenomenal consciousness in any system. Including in our two robots, we could apply phi to it. In fact, we can work that out mathematically some rigor to show that PHI ought to give a different readout for higher consciousness in the case of the robot that is integrated, that uses integrated information.

But we say that humans stand in a severe gap or on the other side of a severe gap or chasm or canyon from non-human animals. And cognitive consciousness predicts that and says that. And they're forced to go the other way, they end up saying everything is. Now, there might be a higher PHI reading, but everything... I'm looking at my EarPod, I'm looking at my apple iPods right in front of me now. So it is going to end up with a phi reading. It's actually computing right now. I opened it up, it sent me out a green signal saying, "I'm fully charged." Or at least I have power enough to charge your pods. And look, is there something it's like to be my iPods? Goodness gracious. I don't-

Pat Flynn:

The panpsychist might say yes, but I think most of us still want to say no, right? Yeah.

Selmer Bringsjord:

I think so. Now, one thing that I wanted to get into with you, and I don't know how it will accord or not with your plans, is that since we wrote the paper, we and they and many, many others have started to explore the quantum case. So that is quantum computing, quantum computation. Because we now have quantum computers. We had the amazing chapter in American corporate battles between Google and IBM that some may be unaware of where Google laid a trap for IBM.

They got a quantum computer and they declared that it had done some amazing things that were unprecedented. They said it has reached quantum supremacy. And then IBM said, "Well, come on. You can't just make that announcement. You have to write a technical paper and you have to disclose it and someone has to referee it." And then the trap was they actually had done that and they announced the paper was coming out. But people have talked about this connection between quantum mechanics and quantum effects in the world of the quantum and consciousness forever really. Roger Penrose, great experience. I had a dinner with him years ago and I think he was just starting to write about these things. And he said, "You know, I do think there might be something in the brain at the quantum level that's relevant." And when he was using the term consciousness, he was talking about phenomenal consciousness.

So these folks are pushing the envelope since the release of our paper, there was no opportunity to write about it anyway. It wouldn't have made sense. But they're pushing the envelope and using... Your listeners can take a look. They have a 2024 paper Koch's the last author, but I think the driver of it, Christoph Koch, they're exploring phenomenal consciousness in a quantum computer and whether that's possible or not. We have no publications. So we've started to look at this from the standpoint of cognitive consciousness, which obviously I have to explain.

But their stuff in this regard, this is really scary stuff. Well, for starters, back to the Apple pods, if they really believe that my apple iPods have some degree of consciousness in the phenomenal sense, I've already parted ways. But let's just take note of the empirical fact, Pat, that they actually believe these things. Naveen and I learned that firsthand in person. These are not people trying to get attention or accolades or anything like that. In fact, many people in AI looking at consciousness have accused them of playing fast and loose in various ways in their publications. There's really severe intellectual battles going on.

Pat Flynn:

Yes.

Selmer Bringsjord:

They know they believe this stuff. Well then they really believe that they're poking around now in the realm of, because of the quantum, the move to quantum computation that we might actually be seeing glimmers of phenomenal consciousness when we are using a quantum computer. Now, let's think about this. Google fired Blake Lemoine I think for saying that his chatbot, in his opinion was conscious. He used the term sentience, but what he meant was that it is subjectively aware that there was something it was like to be it. And he got worried and said, "I don't want any part of this." And they fired him. But wait a minute, the only way you would be justified in firing him is if you provided a precise, compelling rationale for why he was wrong.

Now, I believe he was wrong. They believe he was wrong. It's generally unnoticed. I could say that there are some people who certainly notice it and are worried about it. But just thinking about the sort of basics of Western jurisprudence here, he was convinced that the chatbot had inner feelings, emotions, and that there was something it was like to be it. All right, he's wrong. We could spend an episode or

another episode demonstrating why he's wrong. But now let's think about what these folks are doing in the case of IIT and Phi, in moving to the quantum realm. And it's open access to the entropy paper. If people are skeptical, they don't believe that this is a detailed investigation and that they're really pushing these limits in an amazingly careful way physically, they can go get the article.

That's morally questionable. If they're a hundred percent convinced that this is all it takes, this integrated information happening in a system and they firmly believe that when they move to the quantum level, they'll really have it. From an ethical standpoint, how could it be okay to possibly create exactly what Lemoine thought erroneously was going on in his case? This is really weird stuff. I wasn't a referee, but I would say at least you've got to say something about this issue.

Pat Flynn:

Yes. Right, obviously. Yeah, it's a great point. There's obviously major potential ethical implications if they are correct about this position.

Selmer Bringsjord:

Right. And it's up to God only, I suppose, to judge people for purely mental acts in terms of law. I got that. But we introduce mental states as necessary components of crimes that are physically determinate, right? That's why we have essentially shades of murder, right?

Pat Flynn:

Yep.

Selmer Bringsjord:

And then we go to manslaughter, et cetera. We do take account mental states. I'm not saying what they're doing definitively is morally wrong. I'm saying it sure looks like if that's the beliefs they have, that this could very well be morally wrong. Because what's your upside to this? Especially, we got to talk about cognitive consciousness, but when we can get the gains in the intelligence of a system by looking at something, which since my interactions with John McCarthy on this concept are, "No, there's no consciousness here in that phenomenal sense." I'm using the term to refer to intense levels of consciousness in a structural sense that produces intelligent behavior in the system. "I don't think there's any feelings going on and there'd be no reason for that." Is what McCarthy would say. So I'm just throwing that out there. Naveen and I, we're going to take LaMDA into the quantum realm and see what we can do.

Austin Egbert:

That's all for this week. We'll be back soon. To wrap up this discussion of integrated information theory here at Mind Matters News. Until then, be of good cheer.

Announcer:

This has been Mind Matters News, explore more at Mindmatters.ai. That's Mindmatters.ai. Mind Matters News is directed and edited by Austin Egbert. The opinions expressed on this program are solely those of the speakers.