

SDS PODCAST

EPISODE 748

FIVE-MINUTE

FRIDAY:

THE FIVE LEVELS OF

AGI



(00:00):

This is Five-Minute Friday with The Five Levels of AGI.

(00:27):

Welcome back to the SuperDataScience Podcast. I'm your host, Jon Krohn. So today's episode is all about a paper out of Google DeepMind recently called Levels of AGI: Operationalizing Progress on the Path to AGI. So this paper is all about artificial general intelligence, which is the concept of an algorithm that could have all the general capabilities of a human. But the problem with this AGI definition is that it's really vague. So in Episode #729, for example, professor Blake Richards from the University of Montreal was on the show. He's an expert in neuroscience and AI, and he doesn't believe in the concept of AGI at all because he thinks it's so poorly defined. So these authors from Google DeepMind, DeepMind being one of the premier AI research labs, they have long had the explicit mission of solving artificial general intelligence.

(01:32):

Yeah, so they're interested in being able to define it well. In this paper, again, it's called Operationalizing Progress on the Path to AGI. Of course, we have a link to that paper which is available in arXiv in the show notes for you to download. The final author on this paper is Shane Legg, who is one of the co-founders of AGI. So yeah, AGI clearly on the mind for this AGI-focused company and its co-founding team. So in this paper, they come up with specific levels of AGI, and so you can see all of this in Table 1 of the paper. So there's five levels of AGI in the paper, but there's also a level 0, which is no AI at all. So at that level 0 where there's no AI, an example of a narrow non-AI kind of computing system could be something like the software on a calculator.

(02:28):

In terms of a general kind of non-AI computing capability, so something that is capable on a wide range of non-physical tasks but doesn't involve AI is something like human in-the-loop computing, so like Amazon's Mechanical Turk where there's computing that happens, but you require a human to be there in the loop in order for there to be a wide range of non-physical tasks, including things like learning new skills to be possible in the system. So that's the no AI level, level 0. Level 1 is where we start to actually have AI. So they call level 1 the emerging level. At this level, the AI system has capabilities that are equal to or somewhat better than an unskilled human. So an example of a narrow emerging AI system, so an AI system that is equal to or somewhat better than an unskilled human, but only on a specific task or a narrow set of tasks. An example of that comes from 1971, this algorithm called SHRDLU, which was an early natural language understanding algorithm.

(03:45):

So it was kind of okay at understanding natural language, but that's all it could do. It was only just kind of okay. It was only just about as good as an unskilled human at this very narrow task. Now in terms of emerging AGI, so in terms of an emerging general algorithm, so we're still at this level 1, whereas things are just equal to or somewhat better than an unskilled human, but when we jump from a narrow AI, we're talking about things that have been around 50 years in some cases like that SHRDLU algorithm I was just talking about. Whereas, when we're talking about emerging AGI, we're talking about the absolute cutting edge of things happening today. So things like OpenAI's GPT-4, Google's Bard algorithm or more recently, their Gemini algorithm, Meta's Llama 2 these are all examples of emerging AGI systems. So they're capable on a very wide range of non-physical tasks, making them general. However, they're only just equal to or somewhat better than an unskilled human. All right, so that's level 1, emerging.



(04:48):

The next level is level 2, and this is competent AI. The authors from DeepMind define competent AI as being better than at least 50% of skilled adults, so at least the 50th percentile of skilled adults. So examples of competent narrow AI systems include things like Apple's Siri, Amazon's Alexa, Google Assistant, IBM's Watson, or even state-of-the-art LLMs like GPT-4, Gemini, Llama 2. But those state-of-the-art LLMs, they're only competent at this level of at least 50th percentile of skilled adults on a narrow set of tasks. So that's why this is just competent narrow AI. So the same kinds of large language models, they count as emerging AGI, so at level 1. But at level 2, with this higher threshold of 50th percentile of skilled adults, they're only narrow AI. They're only competent because they only exceed that threshold of half of skilled adults on a very narrow range of tasks, things like short essay writing or simple coding tasks.

(06:06):

So today, as of recording this in January 2024, there are no examples of AGI systems at this competent level that is above half of skilled adults. But nevertheless, moving on, the whole point of this is that we're able to be cataloging the AGI systems that are coming, not just the ones that we already have today. So moving along from the competent level to the expert level, level 3. At level 3, this is an algorithm that is at the level of at least the 90th percentile of skilled adults. So level 2 was 50 percentile, level 3 is 90th percentile. So now we're talking about an AI system that's better than 90% of skilled adults or more. We have lots of examples of expert narrow AI systems today. So Grammarly, which is a spelling and grammar checker, the popular one. Image generation models like DALL·E 2 or Midjourney, those are all examples of AI systems that are, they create stunning art, for example, or they're amazing at checking spelling and grammar.



(07:13):

They're better than 90% of skilled adults at spelling and grammar checking or generating images. But of course, at this expert level, we don't have an AGI system that is widely capable yet. All right, and then onto level 4, this is what the DeepMind authors call the virtuoso level. So this is where the algorithm is at least the 99th percentile of skilled adults. So examples of Virtuoso narrow AI systems include things like IBM's Deep Blue that could beat Garry Kasparov a couple of decades ago at chess. So the world's best chess player being beaten by a chess system, clearly this is a virtuoso narrow AI. It can absolutely crush even the best chess players in the world sometimes. So more recently, an example would be AlphaGo, which also came out of DeepMind. It could beat some of the world's best Go players.

(08:10):

Again, while we have Virtuoso narrow AI systems, of course, we don't have any Virtuoso AGI systems at this level 4, just as we didn't have AGI systems at level 3 or level 2. Then finally, the top level, level 5, this is superhuman performance. So at level 5, the AI system outperforms 100% of humans. So we do actually have superhuman narrow AI systems around today. These are relatively new, but things like AlphaFold, an algorithm out of Google DeepMind themselves, this AlphaFold algorithm is able to very accurately, in many instances, predict the 3D protein structure of a protein based on the two, well, the one-dimensional sequence of amino acids, these building blocks of proteins.

(09:00):

So you have these one-dimensional sequences that are made up of these building blocks called amino acids, and it is a nigh on impossible for human to be able to figure out how that one-dimensional array of amino acids is going to fold up into a working protein in your body. But the AlphaFold algorithm can do that. So it is superhuman, it is, outperforms



100% of humans. If we can create an algorithm that is able to be able to outperform 100% of humans on a wide range of non-physical tasks, including being able to learn new skills, that would be artificial super intelligence or ASI. Yeah, obviously, we haven't achieved that. We don't even know for sure that it is achievable, but it might be, and it might be in our lifetimes. This could be coming in coming decades or at the pace that AI is moving, it could be even sooner potentially. So very exciting times, a very exciting industry to be working in.

(10:06):

I hope you find all of this more exciting than scary. There certainly could be some risks. There are lots of risks associated with the development of these amazingly powerful AGI systems. But yeah, hopefully, we can overcome the risks and things like ASI could help us realize a tremendously positive world, a peaceful world and a world where everyone's fed and don't have to worry about violence and are well-educated and all these kinds of wonderful things that allow humans to flourish. So fingers crossed, and maybe even you can play a part in that.

(10:43):

All right, and that's it for today. I hope today's episode provided you some clarity on different types of AI and where we are with AGI today, specifically at the emerging AGI level 1 level, or the competent narrow AI, level 2 level, depending on how you look at it. Yeah, until next time, keep on rocking it out there, and I'm looking forward to enjoying another round of the Super Data Science Podcast with you very soon.