When Will Robots Gain ‘Common Sense’?

Published 13 September 2022

Christopher Gannatti, CFA
Global Head of Research

2022 has been a year where we have heard different applications of artificial intelligence continually increasing different types of capabilities:

- Large-language-models, exemplified by GPT-31, have become larger and have pointed their capabilities toward more and more areas, like computer programming languages.
- DeepMind further expanded its AlphaFold toolkit, showing predictions of the structure of more than 200 million proteins and making these predictions available at no charge for researchers2.
- There has even been expansion in what’s termed ‘autoML’3 , which refers to low-code machine learning tools that could give more people, without data science or computer science expertise, access to machine learning4.

However, even if we can agree that advances are happening machines are still primarily helpful in discrete tasks and would not possess much in the way of flexibility to respond to many different changing situations in short order.

Intersection of large language models and robotics

Large language models are interesting in many cases for their emergent properties. These giant models may have hundreds of billions if not trillions of parameters. One output could be written text. Another could be something akin to ‘autocomplete’ in coding applications.
But what if you told a robot something like, ‘I’m hungry.’

As a human being, if we hear someone say, ‘I’m hungry’, we can intuit many different things quite quickly based on our surroundings. At a certain time of day, maybe we start thinking of going to restaurants. Maybe we get the smartphone out and think about takeout or delivery. Maybe we start preparing a meal.

A robot would not necessarily have any of this ‘situational awareness’ if it wasn’t fully programmed in ahead of time. We would naturally tend to think of robots as able to perform their specific, pre-programmed functions within the guidelines of precise tasks. Maybe we would think a robot could respond to a series of very simple instructions—telling it where to go with certain key words, what to do with certain additional key words.

‘I’m hungry’—a two-word command with no inherent instructions would be assumed to be impossible.

Google's Pathways Language Model (PaLM) — a start to more complex human/robot interactions

Google researchers were able to demonstrate a robot able to respond, within a closed environment admittedly, to the statement 'I'm hungry.' It was able to locate food, grasp it, and then offer it to the human5.

Google's PaLM model was underlying the robot's capability to take the inputs of language and translate them into actions. PaLM is notable in that it builds in the capability to explain, in natural language, how it reaches certain conclusions6.

As is often the case, the most dynamic outcomes tend to come when one can mix different ways of learning to lead to greater capabilities. Of course, PaLM, by itself, cannot automatically inform a robot how to physically grab a bar of chocolate, for example. The researchers would demonstrate, via remote control, how to do certain things. PaLM was helpful in allowing the robot to connect these concrete, learned actions, with relatively abstract statements made by humans, such as 'I'm hungry' which doesn't necessarily have any explicit command7.

The researchers at Google and Everyday Robots titled a paper 'Do As I Can, Not As I Say: Grounding Language in Robotic Affordance's 8. In Figure 1, we see the genius behind such a title, in that it's important to recognise that large language models may take as their 'inspiration' text from across the entire internet, most of which would not be applicable to a particular robot in a particular situation. The system must 'find the intersection' between what the language model indicates makes sense to do and what the robot itself can actually achieve in the physical world. For instance:

- Different language models might associate cleaning up a spill with all different types of cleaning—they may not be able to use their immense training to see that a vacuum may not be the best way to clean up a liquid. They may also simply express regret that a spill has occurred.

- If one is thinking of the intersection that has the highest chance of making sense, if a robot in a given situation can 'find a sponge' and the large language model indicates that the response of 'find a sponge' could make sense, marrying these two concepts could lead the robot to at least attempt a productive, corrective action to the spill situation.

The 'SayCan' model, while certainly not perfect and not a substitute for true understanding, is an interesting way to get robots to do things that could make sense in a situation without being directly programmed to respond to a statement in that precise manner.



Source: Graphic from Ahn et al. "Do As I Can, Not As I Say: Grounding Language in Robotic Affordances." ARXIV. Submitted 4 April 2022, last revised 16 August 2022.

In a sense, this is the most exciting part of this particular line of research:

- Robots tend to need short, hard-coded commands. Understanding more less specific instructions isn't typically possible.
- Large language models have demonstrated an impressive capability to respond to different prompts, but it's always in a 'digital-only' setting.

If the strength of robots in the physical world can be married with the, at least seeming, capability to understand natural language that comes from large language models, you have the opportunity for a notable synergy that is better than either working on its own.

Conclusion: companies are pursuing robotic capabilities in a variety of ways

Within artificial intelligence, it's important to recognise the critical progression from concept to research to breakthroughs and then only later mass market usage and (hopefully) profitability. The robots understanding abstract natural language today could be some distance away from mass market revenue generating activity.

Yet, we see companies taking action toward greater and greater usage of robotics. Amazon is often in focus for what it may be able to use robots for in its distribution centres, but even more recently it has announced its intention to acquire iRobot9, the maker of the Roomba vacuum system. Robots with increasingly advanced capability will have a role to play in society as we keep moving forward.

Today's environment of rising wage pressures does have companies exploring more and more what robots and automation could bring to their operations. It is important not to overstate where we are in 2022—robots are not able to exemplify fully human behaviours at this point—but we should expect remarkable progress

in the coming years.

Sources

1 Generative Pre-trained Transformer 3
2 Source: Callaway, Ewen. "'The Entire Protein Universe': AI Predicts Shape of Nearly Every Known Protein." Nature. Volume 608. 4 August 2022.
3 Automated machine learning
4 Source: Xu, Tammy. "Automated techniques could make it easier to develop AI." MIT Technology Review. 5 August 2022.
5 Source: Knight, Will. "Google's New Robot Learned to Take Orders by Scraping the Web." WIRED. 16 August 2022.
6 Source: Knight, 16 August 2022.
7 Source: Knight, 16 August 2022.
8 Source: Ahn et al. "Do As I Can, Not As I Say: Grounding Language in Robotic Affordances." ARXIV. Submitted 4 April 2022, last revised 16 August 2022.
9 Source: Hart, Connor. "Amazon Buying Roomba Maker iRobot for $1.7 Billion." Wall Street Journal. 5 August 2022.

Important Risks Related to this Article

Important Information

Marketing communications issued in the European Economic Area ("EEA"): This document has been issued and approved by WisdomTree Ireland Limited, which is authorised and regulated by the Central Bank of Ireland.

Marketing communications issued in jurisdictions outside of the EEA: This document has been issued and approved by WisdomTree UK Limited, which is authorised and regulated by the United Kingdom Financial Conduct Authority.

WisdomTree Ireland Limited and WisdomTree UK Limited are each referred to as "WisdomTree" (as applicable). Our Conflicts of Interest Policy and Inventory are available on request.

For professional clients only. The information contained in this document is for your general information only and is neither an offer for sale nor a solicitation of an offer to buy securities or shares. This document should not be used as the basis for any investment decision. Investments may go up or down in value and you may lose some or all of the amount invested. Past performance is not necessarily a guide to future performance. Any decision to invest should be based on the information contained in the appropriate prospectus and affer seeking independent investment, tax and legal advice.

The application of regulations and tax laws can offen lead to a number of different interpretations. Any views or opinions expressed in this communication represent the views of WisdomTree and should not be construed as regulatory, tax or legal advice. WisdomTree makes no warranty or representation as to the accuracy of any of the views or opinions expressed in this communication. Any decision to invest should be based on the information contained in the appropriate prospectus and affer seeking independent investment, tax and legal advice.

This document is not, and under no circumstances is to be construed as, an advertisement or any other step in furtherance of a public offering of shares or securities in the United States or any province or territory thereof. Neither this document nor any copy hereof should be taken, transmitted or distributed (directly or indirectly) into the United States.

Although WisdomTree endeavours to ensure the accuracy of the content in this document, WisdomTree does not warrant or guarantee its accuracy or correctness. Where WisdomTree has expressed its own opinions related to product or market activity, these views may change. Neither WisdomTree, nor any affiliate, nor any of their respective officers, directors, partners, or employees accepts any liability whatsoever for any direct or consequential loss arising from any use of this document or its contents.