

When Will Robots Gain “Common Sense”?

Published September 12, 2022

Christopher Gannatti, CFA

Global Head of Research

In 2022, we have heard of different applications of artificial intelligence continually increasing different types of capabilities:

- Large language models, exemplified by GPT-3, have gotten 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 to researchers at no charge.¹
- There has even been expansion in what's termed “autoML,” which refers to low-code machine learning tools that could give more people, without data science or computer science expertise, access to machine learning.²

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

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, we may start thinking of going to restaurants. Maybe we get the smartphone out and think about takeout or delivery. Maybe we start preparing that next meal for our family.

A robot would not necessarily have any of this ‘situational awareness’ if it wasn’t fully programmed ahead of time. We 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” is a two-word command with no inherent instructions that would be assumed to be impossible for a robot to respond to appropriately.

Google's PaLM Language Model—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 offer it to the person.³

Google's PaLM model supported the robot's ability 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 conclusions.⁴

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 bag of chips, for example. The researchers demonstrated via remote control how to do certain things. PaLM was helpful then in allowing the robot to connect these concrete, learned actions with relatively abstract statements made by humans, such as “I’m hungry,” that don't necessarily have any explicit command.⁵

The researchers at Google and Everyday Robots wrote a paper, “Do as I Can, Not as I Say: Grounding Language in Robotic Affordances.”⁶ In figure 1, we see the genius behind such a title, in that it's important to recognize 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—but 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 a robot in a given situation can ‘find a sponge’ and the large language model indicates that the response ‘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.

Figure 1: Illustrative Representation of How ‘SayCan’ Could Work



Source: Ahn et al, "Do as I Can, Not as I Say: Grounding Language in Robotic Affordances," ARXIV, submitted 4/4/22, last revised 8/16/22.

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

- Robots tend to need short, hard-coded commands. Understanding 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 apparent 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 recognize the critical progression from concept to research to breakthroughs and then only later to 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 centers, but even more recently, it has announced its intention to acquire iRobot, 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 behaviors at this point—we believe we will see remarkable progress in the coming years.

Investors looking to gain an exposure to a diverse array of companies focused on artificial intelligence should consider the [WisdomTree Artificial Intelligence and Innovation Fund \(WTAI\)](#).

1 Source: Ewen Callaway, Ewen. "'The Entire Protein Universe': AI Predicts Shape of Nearly Every Known Protein," *Nature*, 8/4/22.

2 Source: Tammy Xu, "Automated Techniques Could Make it Easier to Develop AI," *MIT Technology Review*, 8/5/22.

3 Source: Will Knight, "Google's New Robot Learned to Take Orders by Scraping the Web," *WIRED*, 8/16/22.

4 Source: Knight, 16 August 2022.

5 Source: Knight, 16 August 2022.

6 Source: Ahn et al, "Do as I Can, Not as I Say: Grounding Language in Robotic Affordances," *ARXIV*, submitted 4/4/22, last revised 8/16/22.

Christopher Gannatti is an employee of WisdomTree UK Limited, a European subsidiary of WisdomTree Asset Management, Inc.'s parent company, WisdomTree Investments, Inc.

As of September 9, 2022, WTAI held 0%, 0%, 1.31% and 0% in Google, Everyday Robots, Amazon and iRobot, respectively. DeepMind is a subsidiary of Alphabet. As of September 9, 2022, Alphabet was a 1.37% exposure in WTAI.

Important Risks Related to this Article

There are risks associated with investing, including the possible loss of principal. The Fund invests in companies primarily involved in the investment theme of artificial intelligence (AI) and innovation. Companies engaged in AI typically face intense competition and potentially rapid product obsolescence. These companies are also heavily dependent on intellectual property rights and may be adversely affected by loss or impairment of those rights. Additionally, AI companies typically invest significant amounts of spending on research and development, and there is no guarantee that the products or services produced by these companies will be successful. Companies that are capitalizing on Innovation and developing technologies to displace older technologies or create new markets may not be successful. The Fund invests in the securities included in, or representative of, its Index regardless of their investment merit and the Fund does not attempt to outperform its Index or take defensive positions in declining markets. The composition of the Index is governed by an Index Committee and the Index may not perform as intended. Please read the Fund's prospectus for specific details regarding the Fund's risk profile.