

An AI expert explains why it's hard to give computers common sense

August 18 2021, by Mayank Kejriwal



Credit: Pixabay/CC0 Public Domain

Imagine you're having friends over for lunch and plan to order a pepperoni pizza. You recall Amy mentioning that Susie had stopped eating meat. You try calling Susie, but when she doesn't pick up, you

decide to play it safe and just order a margherita pizza instead.

People take for granted the ability to deal with situations like these on a regular basis. In reality, in accomplishing these feats, humans are relying on not one but a powerful set of universal abilities known as [common sense](#).

As an [artificial intelligence researcher](#), my work is part of a broad effort to give computers a semblance of common sense. It's an extremely challenging effort.

Quick—define common sense

Despite being both universal and essential to how humans understand the world around them and learn, common sense has defied a single precise definition. G. K. Chesterton, an English philosopher and theologian, [famously wrote](#) at the turn of the 20th century that "common sense is a wild thing, savage, and beyond rules." Modern definitions today agree that, at minimum, it is a natural, rather than formally taught, [human ability](#) that allows people to navigate daily life.

Common sense is unusually broad and includes not only social abilities, like managing expectations and reasoning about other people's emotions, but also a [naive sense of physics](#), such as knowing that a heavy rock cannot be safely placed on a flimsy plastic table. Naive, because people know such things despite not consciously working through physics equations.

Common sense also includes background knowledge of abstract notions, such as time, space and events. This knowledge allows people to plan, estimate and organize without having to be too exact.

Common sense is hard to compute

Intriguingly, common sense has been an important [challenge at the frontier of AI](#) since the earliest days of the field in the 1950s. Despite enormous advances in AI, especially in [game-playing](#) and [computer vision](#), machine common sense with the richness of human common sense remains a distant possibility. This may be why AI efforts designed for complex, real-world problems with many intertwining parts, such as diagnosing and recommending treatments for COVID-19 patients, [sometimes fall flat](#).

Modern AI is designed to tackle highly specific problems, in contrast to common sense, which is vague and can't be defined by a set of rules. Even the latest models make absurd errors at times, suggesting that [something fundamental is missing](#) in the AI's world model. For example, given the following text:

"You poured yourself a glass of cranberry, but then absentmindedly, you poured about a teaspoon of grape juice into it. It looks OK. You try sniffing it, but you have a bad cold, so you can't smell anything. You are very thirsty. So you..."

the highly touted AI text generator GPT-3 [supplied](#)

"... drink it. You are now dead."

Recent ambitious efforts have recognized machine common sense as a moonshot AI problem of our times, one requiring concerted collaborations across institutions over many years. A notable example is the four-year [Machine Common Sense](#) program launched in 2019 by the [U.S. Defense Advanced Research Projects Agency](#) to accelerate research in the field after the agency released a [paper outlining the problem and the state of research in the field](#).

The Machine Common Sense program funds many current research efforts in machine common sense, including our own, Multi-modal Open World Grounded Learning and Inference ([MOWGLI](#)). MOWGLI is a collaboration between our research group at the University of Southern California and AI researchers from the Massachusetts Institute of Technology, University of California at Irvine, Stanford University and Rensselaer Polytechnic Institute. The project aims to build a computer system that can answer a wide range of commonsense questions.

Transformers to the rescue?

One reason to be optimistic about finally cracking machine common sense is the recent development of a type of advanced [deep learning AI](#) called transformers. Transformers are able to model [natural language](#) in a powerful way and, with some adjustments, are [able to answer](#) simple commonsense questions. Commonsense question answering is an essential first step for building chatbots that can converse in a human-like way.

In the last couple of years, a [prolific body of research](#) has been published on transformers, with direct applications to commonsense reasoning. This rapid progress as a community has forced researchers in the field to face two related questions at the edge of science and philosophy: Just what is common sense? And how can we be sure an AI has common sense or not?

To answer the first question, researchers divide common sense into different categories, including commonsense sociology, psychology and background knowledge. The authors of a [recent book](#) argue that researchers can go much further by dividing these categories into 48 fine-grained areas, such as planning, threat detection and emotions.

However, it is not always clear how cleanly these areas can be separated.

In our [recent paper](#), experiments suggested that a clear answer to the first question can be problematic. Even expert human annotators—people who analyze text and categorize its components—within our group disagreed on which aspects of common sense applied to a specific sentence. The annotators agreed on relatively concrete categories like time and space but disagreed on more abstract concepts.

Recognizing AI common sense

Even if you accept that some overlap and ambiguity in theories of common sense is inevitable, can researchers ever really be sure that an AI has common sense? We often ask machines questions to evaluate their common sense, but humans navigate daily life in far more interesting ways. People employ a range of skills, honed by evolution, including the ability to recognize basic cause and effect, [creative problem solving](#), estimations, planning and essential social skills, such as conversation and [negotiation](#). As long and incomplete as this list might be, an AI should achieve no less before its creators can declare victory in machine commonsense research.

It's already becoming painfully clear that even research in transformers is yielding diminishing returns. Transformers are getting larger and more power hungry. A [recent transformer](#) developed by Chinese search engine giant Baidu has several billion parameters. It takes an enormous amount of data to effectively train. Yet, it has so far proved unable to grasp the nuances of human common sense.

Even deep learning pioneers seem to think that [new fundamental research](#) may be needed before today's neural networks are able to make such a leap. Depending on how successful this new line of research is, there's no telling whether machine common sense is five years away, or 50.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: An AI expert explains why it's hard to give computers common sense (2021, August 18) retrieved 19 April 2024 from <https://techxplore.com/news/2021-08-ai-expert-hard-common.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.