Infants outperform AI in 'commonsense psychology'

February 21 2023

Credit: Unsplash/CC0 Public Domain

Infants outperform artificial intelligence in detecting what motivates other people's actions, finds a new study by a team of psychology and data science researchers. Its results, which highlight fundamental

differences between cognition and computation, point to shortcomings in today's technologies and where improvements are needed for AI to more fully replicate human behavior.

"Adults and even infants can easily make reliable inferences about what drives other people's actions," explains Moira Dillon, an assistant professor in New York University's Department of Psychology and the senior author of the paper, which appears in the journal *Cognition*. "Current AI finds these inferences challenging to make."

"The novel idea of putting infants and AI head-to-head on the same tasks is allowing researchers to better describe infants' intuitive knowledge about other people and suggest ways of integrating that knowledge into AI," she adds.

"If AI aims to build flexible, commonsense thinkers like human adults become, then machines should draw upon the same core abilities infants possess in detecting goals and preferences," says Brenden Lake, an assistant professor in NYU's Center for Data Science and Department of Psychology and one of the paper's authors.

It's been well-established that infants are fascinated by other people—as evidenced by how long they look at others to observe their actions and to engage with them socially. In addition, previous studies focused on infants' "commonsense psychology"—their understanding of the intentions, goals, preferences, and rationality underlying others' actions—have indicated that infants are able to attribute goals to others and expect others to pursue goals rationally and efficiently. The ability to make these predictions is foundational to human social intelligence.

Conversely, "commonsense AI"—driven by machine-learning algorithms —predicts actions directly. This is why, for example, an ad touting San Francisco as a travel destination pops up on your computer screen after

you read a news story on a newly elected city official. However, what AI lacks is flexibility in recognizing different contexts and situations that guide human behavior.

To develop a foundational understanding of the differences between humans' and AI's abilities, the researchers conducted a series of experiments with 11-month-old infants and compared their responses to those yielded by state-of-the-art learning-driven neural-network models.

To do so, they deployed the previously established "Baby Intuitions Benchmark" (BIB)—six tasks probing commonsense psychology. BIB was designed to allow for testing both infant and machine intelligence, allowing for a comparison of performance between infants and machines and, significantly, providing an empirical foundation for building human-like AI.

Specifically, infants on Zoom watched a series of videos of simple animated shapes moving around the screen—similar to a video game. The shapes' actions simulated human behavior and decision-making through the retrieval of objects on the screen and other movements. Similarly, the researchers built and trained learning-driven neural-network models—AI tools that help computers recognize patterns and simulate human intelligence—and tested the models' responses to the exact same videos.

Their results showed that infants recognize human-like motivations even in the simplified actions of animated shapes. Infants predict that these actions are driven by hidden but consistent goals—for example, the on-screen retrieval of the same object no matter what location it's in and the movement of that shape efficiently even when the surrounding environment changes. Infants demonstrate such predictions through their longer looking to such events that violate their predictions—a common and decades-old measurement for gauging the nature of infants'

knowledge. Adopting this "surprise paradigm" to study machine intelligence allows for direct comparisons between an algorithm's quantitative measure of surprise and a well-established human psychological measure of surprise—infants' looking time. The models showed no such evidence of understanding the motivations underlying such actions, revealing that they are missing key foundational principles of commonsense psychology that infants possess.

"A human infant's foundational knowledge is limited, abstract, and reflects our evolutionary inheritance, yet it can accommodate any context or culture in which that infant might live and learn," observes Dillon.

The paper's other authors are Gala Stojnić, an NYU postdoctoral fellow at the time of the study, Kanishk Gandhi, an NYU research assistant at the time of the study, and Shannon Yasuda, an NYU doctoral student.

  More information: Gala Stojnić et al, Commonsense psychology in human infants and machines, *Cognition* (2023). DOI: 10.1016/j.cognition.2023.105406

  Provided by New York University