P H I L M A T H   S E M I N A R

Research seminar in the philosophy of mathematics and the philosophy of logic

Organizers: Fabrice Pataut (SND) and Francesca Poggiolesi (IHPST)

1st SEMESTER 2024-2025 PROGRAM
IHPST – salle de conférences (2ème étage)
13, rue du Four, 75006 Paris

14 Octobre 2024, 17h-19h

Wesley H. Holliday
(University of California, Berkeley)

https://philosophy.berkeley.edu/holliday

Conditional and Modal Reasoning in Large Language Models

The reasoning abilities of large language models (LLMs) are the topic of a growing body of research in AI and cognitive science. In this paper, we probe the extent to which twenty-five LLMs are able to distinguish logically correct inferences from logically fallacious ones. We focus on inference patterns involving conditionals (e.g., "If Ann has a queen, then Bob has a jack") and epistemic modals (e.g., "Ann might have an ace", "Bob must have a king"). These inferences have been of special interest to logicians, philosophers, and linguists, since they play a central role in the fundamental human ability to reason about distal possibilities. Assessing LLMs on these inferences is thus highly relevant to the question of how much the reasoning abilities of LLMs match those of humans. Among the LLMs we tested, all but the GPT-4 model family often make basic mistakes with conditionals, though zero-shot chain-of-thought prompting helps them make fewer mistakes. Moreover, even the GPT-4 family displays logically inconsistent judgments across inference patterns involving epistemic modals, and almost all models give answers to certain complex conditional inferences widely discussed in the literature that do not match human judgments. These results highlight gaps in basic logical reasoning in today's LLMs.

Lien zoom

https://pantheonsorbonne.zoom.us/j/92976738762?pwd=sMF5QppvdGxzatZt5viNllGBUb1L84.1

ID réunion : 929 7673 8762
Code d'accès : 130866