

Thinking Fast and Slow in AI: A Cognitive Architecture to Augment Both AI and Human Reasoning

Francesca Rossi ✉

IBM Research, Yorktown Heights, NY, USA

Abstract

AI systems are very useful in practically every sector, but they also have several limitations, mostly related to the lack of reasoning capabilities. According to the fast and slow thinking theory of human decision making, we can say that data-driven AI, including generative AI, are providing fast thinking capabilities, but they do not have slow thinking ones. Existing cognitive theories of human decision making, such as the thinking fast and slow theory, can provide insights on how to advance AI systems towards some of these capabilities. In this talk I will present a general architecture, called SOFAI, that is based on fast/slow solvers and a meta-cognitive component that provides a centralized governance of the solvers. I will describe two instances of this architecture, for constrained grid navigation and planning, showing experimentally that SOFAI generates better decisions than each of the *individual* solvers. Emerging behavior related to adaptability, skill learning, and cognitive control are also showed in the analysis of SOFAI's behavior. I will also describe how the thinking fast and slow theory can help design a value-based human-machine collaborative decision environment.

2012 ACM Subject Classification Computing methodologies → Cognitive science

Keywords and phrases Artificial Intelligence, Meta-reasoning

Digital Object Identifier 10.4230/LIPICs.CP.2024.2

Category Invited Talk



© Francesca Rossi;

licensed under Creative Commons License CC-BY 4.0

30th International Conference on Principles and Practice of Constraint Programming (CP 2024).

Editor: Paul Shaw; Article No. 2; pp. 2:1–2:1



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany