![]cai{: .align-right} 📖 I'm a final-year M.Eng student at Harvard in Computational Science Engineering, with a keen interest in cognitive AI, computational cognitive/neuroscience, and human-centric AI. Additionally,

🤖 I explore the application of ML and AI in computational social science, particularly in economic and financial markets. Currently, my master's thesis delves into multi-modal models, theory of mind, and deep learning.

🎓 Excitedly, I'm actively seeking a PhD position for Fall 2024. Please get in touch if you think we have

Selected Experience

👓 **Harvard AR/VR Club: Project J[AR]vis** We constructed a highly intelligent AR assistant, J[AR]VIS, merging cognitive architectures and large language models to deliver information automatically or upon user queries. J[AR]VIS allows users to ask questions about their surroundings and receive text responses. My significant contributions include implementing Azure's dense captioning system, processing user questions and scene images for intelligent responses. Additionally, I played a crucial role in developing the cognitive architecture using langchain, deciding which models to use based on image content and enabling various functions, such as internet searches. I also contributed to coding for refining results generated by dense captioning, ensuring accuracy and clarity in the system's responses.

📑 **Do LLMs have theory-of-mind intelligence? We have our own answers...** Since the advent of GPT-4, debates about AGI emergence have intensified due to its impressive capabilities. While some argue that GPT-4 meets critical benchmarks of human intelligence, my exploration suggests it falls short, particularly in the 'Theory of Mind' domain. Large language models (LLMs) like GPT give an AGI impression due to language being an abstraction of human culture, knowledge, and values. However, these models, primarily pattern recognition systems, differ significantly from true AGI. In our research, we introduced the Multimodal Theory-of-Mind Question Answering dataset (MMToM-QA) and developed the BIP-ALM approach, addressing the ToM capabilities gap effectively. Accepted at NeuralPS 2023 @ FMDM and under review for ICLR 2024, our findings advance AI. My involvement deepened my understanding of theory-of-mind, enhanced my deep learning and PyTorch skills, and broadened my perspective on neuro-symbolic AI and zero-shot/few-shot learning, forming the basis of my master's thesis on expanding ToM datasets and models.