Proposal for a 2023 AAAI Fall Symposium on
Integration of Cognitive Architectures and Generative Models
October 25-27, 2023, Washington, D. C.

The purpose of this symposium is to bring together researchers from multiple subdisciplines of AI to explore possible integrations of cognitive architectures and generative models, two very different approaches for developing intelligent agents.

Cognitive Architectures: A major hypothesis of cognitive architecture research has been that human-like minds require the tight integration of multiple, distinct functional components and that those components are on the order of short and long-term memories, decision-making, perception and action, and learning mechanisms. A wide variety of cognitive architectures, including symbolic, neural, and hybrid architectures, have been developed over the years (Kotseruba & Tsotso, 2020). They have often been the basis for creating state-of-the-art models of human behavior as well as complex autonomous AI agents. There have also been efforts to create an abstract consensus across many of these cognitive architectures, called the *Common Model of Cognition*. However, to date, there has been limited research on instilling cognitive architectures with capabilities found in generative models (including massive knowledge). Within the context of the symposium, we will not limit the exploration of architectural organizations to existing cognitive architectures but be open to novel, alternative approaches that provide the overall infrastructure/organization of building intelligent agents.

Generative Models: Recent research in generative models has demonstrated that it is possible to create AI systems through extensive training with massive bodies of pre-existing knowledge across different modalities. Systems such as ChatGPT3 have demonstrated remarkable generality and fluency in language. However, an open question is how to integrate a generative model with other components that are usually assumed to be required when developing general intelligent agents, such as decision-making, memory, online learning, and planning to name a few. Within the context of the symposium, we will consider existing large language models, but also generative models for perception and action (or multiple modalities), as well as how the underlying technologies (such as transformers) can be components or mechanisms of agents.

The goal of the symposium is to explore possible alternative approaches to integrating generative models within a cognitive architecture. Some existing approaches are already being explored in research groups, but we are also interested in hypothetical architectures yet to be implemented. An important goal is to understand the strengths, weaknesses, and tradeoffs among alternative approaches.

Participation:

Our goal is to include a broad range of researchers whose interests span cognitive architecture and generative models. We will require potential participants to provide a one-page description of their interest in the symposium topics. We will also accept longer submissions (details to be determined later) that will be distributed to participants before the symposium and that will be the basis for deciding on presentations as described below.

Schedule:

The symposium will be organized into a series of five sessions over 2.5 days. The exact organization will depend on the attendees and submissions, but our general plan is to have sessions on possible approaches to integration at the architecture level, as well as overlaps in technologies and capabilities between cognitive architectures and generative models. The final morning session on the third day will be a chance to reflect on and discuss future directions. At this point, the potential sessions on the first two days could include:

1. Generative models as components of existing cognitive architectures, such as GPT-4 being a static (but large) knowledge source that is deliberately queried, or a transformer-based perceptual module.
2. Novel architectures where multiple components are generative models connected to memories in a fixed organization, such as in Generative Agents (Park et al. 2024) or stacked LLMs (AutoGPT, BabyAGI?).
3. Commonalities between technologies underlying generative models and cognitive architecture processes – especially hybrids (graphical models, neural networks, …).
4. Explorations/analyses of how different cognitive capabilities, such as decision-making, planning, meta-cognition (including theory of mind), online learning, etc., are or could be achieved in generative models versus cognitive architectures. And conversely, are there mappings of the different aspects of generative models onto the components we associate with existing cognitive architectures?

Each session (three and a half hours including 30 min break) will consist of either an invited talk (30-45min) or an introduction to the topic by the session lead lasting about 10-15 minutes, followed by short (~15 minutes) presentations by (~3-5) participants who have submitted papers related to the session topic. We plan to have significant time devoted to group discussions and potentially break-out groups for the second half of each session.

Contributions to this symposium may present existing or hypothetical examples of architectures and should include discussions of the approach's strengths, weaknesses, tradeoffs, and challenges. We also look for analyses of specific cognitive capabilities and how they can be realized in generative approaches versus cognitive architectures. We will encourage a diverse set of ideas and open discussions.

Organizing Committee:

John Laird (IQMRI), john.laird@cic.iqmri.org,
Center for Integrated Cognition, 311 Windycrest Drive, Ann Arbor, MI 48105 USA
Christian Lebiere (Carnegie Mellon University), cl@cmu.edu,
Psychology Department, Carnegie Mellon University, 5000 Forbes Avenue,
Pittsburgh, PA 15213 USA
Paul S. Rosenbloom (University of Southern California), rosenbloom@usc.edu, Thomas Lord Department of Computer Science.
Andrea Stocco (University of Washington), stocco@uw.edu
Department of Psychology and the Institute for Learning and Brain Sciences, Portage Bay Building, Box 357988, University of Washington, Seattle, WA 98195-7988

We are actively looking for researchers in the Generative Models community to add to the organizing committee.

May 15, 2023