Leveraging Conflict to Bridge Cognitive Reasoning and Generative Algorithms

Anita Raja[1], **Alisa Leshchenko**[1], **Jihie Kim**[2]

[1] Department of Computer Science, Hunter College, City University of New York, NY, NY 10065
[2] Department of Artificial Intelligence, Dongguk University, South Korea
anita.raja@hunter.cuny.edu, alisa.leshchenko25@myhunter.cuny.edu, jihie.kim@dgu.edu

Abstract

Autonomous agents require the ability to identify and adapt to unexpected conditions. Real-world environments are rarely stationary, making it problematic for agents operating in such environments to learn efficient policies. There is therefore a need for a general framework capable of detecting when an agent has encountered novel conditions, and determining how it should adjust its actions. In this position paper we propose a framework that couples cognitive reasoning and generative algorithms by leveraging conflict detection to adjust to unexpected dynamics. Specifically, we propose that a metacognitive conflict resolution mechanism is necessary; such a mechanism would balance the use of commonsense and deliberative reasoning to allow the agent to navigate novel conditions.

Introduction

Consider the following motivating scenarios: Scenario 1: The vision system of a self driving autonomous agent that learns to operate in average weather conditions will find that several of its model assumptions are violated while operating when there is a flash flood caused by an unexpected thunderstorm. However, equipping the agent with the ability to determine that the differences observed (hazy view, washed out roads, high water levels, downed trees, stopped vehicles) are due to change in environmental expectations, would enable it to quickly reframe its assumptions and its policy appropriately. Scenario 2: A caption generation system for an image of West African dining can be affected by the training data that is biased towards images from Western and Eastern cultures, and may have to adapt its expectation of African dining so that the environment and the dishes are properly described.

Learning efficient policies in real-world, resource-bounded environments can be challenging. One reason, as observed in the above example scenarios, is non-stationarity, which occurs when the expectations of goals, location, action outcomes at execution time are no longer the same as those at training time. This non-stationarity can be caused by an agent's own actions, uncertainty in action outcomes, bounded resources, lack of sufficient information during the training phase, the actions of other agents and so on. It often violates the identically distributed assumption of learning algorithms where the training and test data are assumed to be from the same distribution. Out of distribution (OOD) detection and generalization has recently generated significant interest in the machine learning field (Liu et al. 2023). In this position paper, we propose a framework described in Figure 1 that builds on the Common Model of Cognition (CMC) (Laird, Lebiere, and Rosenbloom 2017). It explores a new vein of enquiry to integrate cognitive reasoning with generative model technologies by introducing metalevel knowledge and control to address some of the challenges caused by non-stationarity. The basic premise of this work is that agents operating in non-stationary environments will face conflicts exhibited in the form of incorrect predictions. These conflicts can trigger the need for generating new metalevel knowledge and metalevel control that can non-myopically refocus resources to redo the appropriate context-dependent computations.

While the CMC framework itself is not considered a cognitive architecture, it serves as a repository of collective consensus within the community regarding a unified perspective on numerous cognitive architectures (Laird, Newell, and Rosenbloom 1987; Anderson et al. 2004; Kieras and Meyer 1997). Essentially, it offers an abstract specification of cognitive capabilities, comprised of a series of interconnected modules, with the purpose of abstracting the intricate processes associated with neural processing and complex cognition. CMC supports an expansion of the problem solving process, viz. deliberative reasoning, by adding new relations in the declarative memory or changing the preferences among alternative actions. We propose to augment the model to support out of distribution generalization more explicitly. Specifically, we discuss the use of metalevel control for conflict resolution that bridges cognitive reasoning and generative algorithms

Metalevel Control

Our prior work in single agent (Cox and Raja 2011; Kim et al. 2011) and multiagent (Raja and Lesser 2007; Alexander et al. 2007) metalevel control defines it as the capability of an agent to improve its decision utility by spending some effort to decide what and how much reasoning to do as opposed to what actions to do. It can be viewed as a conductor equipped with a non-myopic view of the agent's problem solving process to ensure greater confidence in its ac-

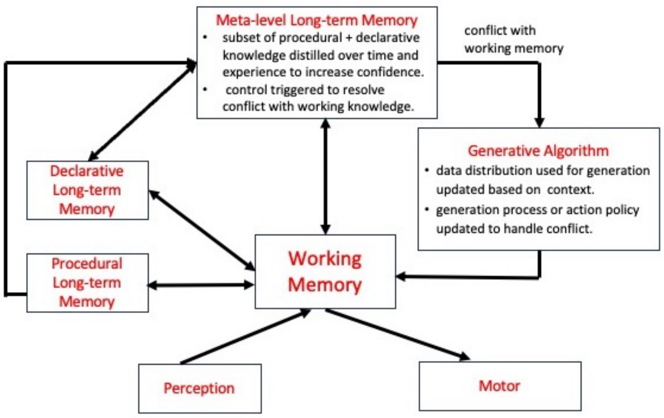

Figure 1: Bridging Deliberative Reasoning and Generative Algorithms via Metalevel Control

tions. Metalevel control, thus, orchestrates when and what knowledge to use, when to trigger cognitive thinking and knowledge and data generative processes and how many resources (time, computation) to invest in each. We define metalevel knowledge as a subset of the agent's declarative and procedural knowledge. This metalevel knowledge gets concretized/distilled in metalevel memory over time such that the certainty and confidence level in the knowledge is above a certain threshold. Metalevel knowledge generally captures social norms; generalized knowledge about world subject to culture, personal experience (family, country); expectation of behavior; knowledge about ones skills, abilities, capacities. It will become a basis for updating the existing knowledge or expanding it. This is related to 'self-regulation' in human learning where we can recognize deficiency of our knowledge based on external feedback and correct or expand them as needed (Bereiter and Scardamalia 2018). The system can measure semantic inconsistencies between the existing knowledge and situation (such as visual information) that it faces based on an assimilation of the knowledge (Chae and Kim 2022). For example, in visual question answering, semantic similarity or dissimilarity can be estimated based on the caption generated from the image and the information retrieved from the knowledge base for the given question. The knowledge can include both the explicit knowledge in the system and the generated knowledge from transformer-based language models. Importantly, metalevel knowledge also contains statistics on what *strategies* have worked well in a given abstract context in the past – a higher order analogue to remembering what actions have worked well in a given state of the environment. Metalevel control is then able to use this knowledge to adaptively guide the agent towards efficient and robust performance even in the context of non-stationary environments.

Role of Conflict: In our framework, conflict is defined as disagreement between an agent's expectations and its observations. Expectations are generated using a combination of implicit (generative models) and explicit (knowledge database) processes. Rather than being an impediment to the agent, conflict is a useful cue that an agent must refine its use of different cognitive strategies and update its policy.

In our proposed framework, metalevel control is triggered when there is a conflict between working knowledge and expectation in agent/environmental state. This conflict can be characterized by the contradiction in expectation of the environmental characteristics in Scenario 1 and social norms in Scenario 2 discussed above, resulting in a need to update of the knowledge or policy about self driving and the distribution of cultural information in generating descriptive sentences associated with the African dining image respectively. In prior work (Cheng, Raja, and Lesser 2013), we designed agents capable of *learning when to learn* new policies in a cooperative multiagent weather tracking environment. We argued that agents with a myopic view of themselves and other agents would at times make inconsistent choices. The decentralized decision makers in the weather tracking application were equipped with a two-phased learning process to determine and obtain the minimal overlapping context required to make their decisions more consistent. This allowed agents to augment their local policy states with metaknowledge of other agents and recompute their policies resulting in more appropriate actions. We posit that this notion of learning when to learn, including when and what context to obtain problem solving context information would be applicable to handling the dynamics caused by non-stationarity.

Out of distribution (OOD) data is data that comes from a distribution other than the one an agent was trained on. OOD conditions are a natural consequence of dynamic environments, and are commonplace in real-world settings. When an agent's observations are determined to be OOD, it can conclude that the environment has changed in a novel way. This can be viewed as a type of conflict. Consider the example of the self-driving car trained in average weather conditions: it is important that the self-driving agent is able to detect that it should not employ the fair weather policy in a flash flood. But what should the agent do? A sensible course of action would be to attempt to resolve its beliefs and to communicate the anomaly to the driver. In our framework, this is handled by metalevel control.

Role of Uncertainty: Furthermore, we believe the challenge for metalevel control is to handle uncertainty at different levels (Alexander et al. 2007). It has to first be able to handle the uncertainty that domain level performance may not always improve by devoting more resources to a deliberative option. Furthermore, it has to handle the uncertainty whether the deliberative decision will be required at all since the execution trajectory may deviate in an unexpected direction. Thus the metalevel policy should be able to consider counterfactual scenarios and determine the utility of alternate deliberative options given possible future state trajectories. Metalevel control should be non-myopic from both a temporal sense (look at future possibilities) but also from a physical sense (have models of itself and also of other agents in its environment).

Role of Resource Bounds According to Kahneman's theory of thinking fast and slow (Kahneman 2011), slow thinking (system 2, deliberative) requires most effort when it has to be done fast, i.e under strict time constraints. He also states that switching between system 1 (intuitive, unconscious decisions) and system 2 and vice versa under time pressure is effortful. Recent work (Booch et al. 2021) has argued for the need for the governance of the two types of thinking and determining when to do the switch and interjecting the notion of time, resources and divergence resolution has been made. In our view, these issues naturally fall under the purview of metalevel control. The use of abstract representation of the agent state and metalevel control with bounded computational overhead has been shown (Raja and Lesser 2007) to result in efficient performance of complex agents in dynamic open multi-agent environments.

Low Level Processing

The above sections have dealt primarily with higher order cognition, such as deliberative reasoning, planning/scheduling, metacognition, and multi-agent coordination. An agent capable of navigating real-world environments will by definition be embodied, and therefore low level perceptual processes are also important to address. In particular, we posit that incoming sense data must be processed before it can be used, as a real-world environment provides raw data rather than pre-computed, meaningful perceptions. Our framework includes a low-level perception cycle (Figure 2) that triggers higher level reasoning processes via attention, only when necessary, thereby saving the agent computational cycles. Note: Here by attention we mean a low-level mechanism capable of interrupting the agent, orienting it to specific information, and placing it into a different mode of functioning.

Perception Cycle

As the agent interacts with the environment, it is necessary for it to continually gather new sense data to test its perceptual beliefs and monitor the results of its actions. In our agent framework, the perceptual cycle will center around prediction and expectation; for the purposes of this discussion, we formulate it in the context of a vision problem. Sense data will be gathered in a *foveated* fashion – meaning high detail resolution in a small region (the 'fovea') and low detail resolution across the rest of the visual field. When sense data is ingested, long-term memory will be queried for information on the current context. This will be done using associative memory, such that partial contextual information gathered from the environment can act as a cue to retrieve relevant long-term memories.

Sense data and contextual information will then be passed to a generative model, which will assist in making a prediction about what the agent should expect to see. In particular, the model will fill in detail in the periphery based on context and will predict the next visual frame. If the agent holds relevant higher-order beliefs, these will also used to inform the prediction. The output of the generative model will be the agent's perception.

The agent then compares its perception at time t with its prediction from time t-1. If the error is too great – that is, if the agent's expectations are violated – the agent attempts to resample the environment. This is analogous to 'taking a closer look'. Assuming sufficient degrees of freedom in

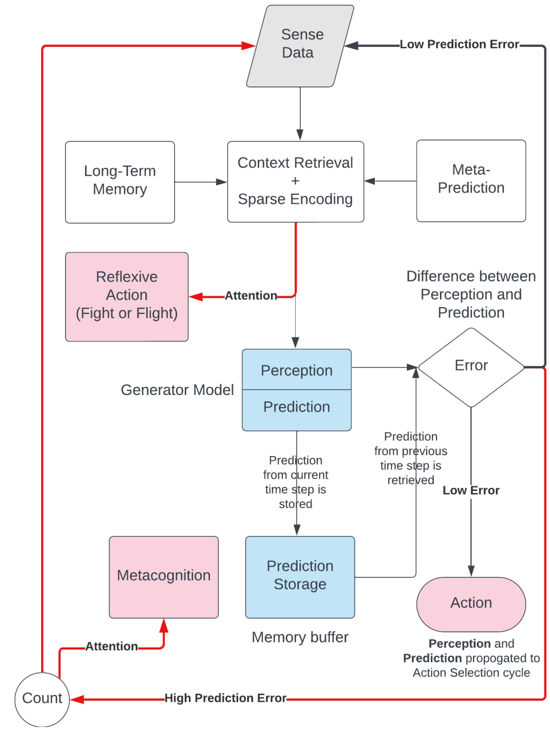

Figure 2: Perception cycle

the agent's control over its body, taking another look will result in visual sense data sampled at a slightly different angle. This small perturbation in the angle of sight should result in a small change in the perceived scene. If the surprising observation passes this stability check, attention is triggered and passes information on to higher cognitive functions. At this higher cognitive level, the agent may employ strategies such as abductive reasoning to find the most likely explanation for its surprising observation. If the stability check is not passed and the surprising element is not encountered upon a second look, then it is safe to assume the perception was likely a mistake. In this case, the agent resumes its normal functioning and goes on to select an action.

Illustrative example: In a dining setting, if the agent mistakenly perceives a meat dish on the table during dessert, its expectations will be violated. It will then take a second look at the surprising region of visual space. From a slightly different angle (and due to the stochastic nature of foveated vision), it is unlikely to make the same mistake again. If, however, it still perceives a meat dish, higher cognition will be triggered to attempt to deal with the situation. First, the agent will search its memory for a likely explanation (abductive reasoning). For instance, it may recall that in some cuisines, sweetened meat is eaten for dessert. If no simple explanation is found, more expensive deliberative reasoning will be triggered. The agent will think deeply about what may have caused the meat dish to be served, and will probably alert others to the situation.

Importantly, the central role of prediction in our framework will bypass several issues. An autonomous agent must have the ability to learn unsupervised (i.e. without access to ground truth). Using predictive error circumvents this issue. In addition, as long as predictions are reliable, they allow for faster reaction time than using solely present observations. Note: Low level aspects of our framework are similar to but distinct from predictive coding(A. 2013).

Bridge to Higher Cognition

In future work, we will further develop the interface between lower-level perceptual processes and higher cognition. This is a complex subject that touches on issues like the relationship between symbols and meaning. Here we will include a brief sketch of the way this will be handled in our framework. Firstly, just as raw sense data must be processed into perceptions, perceptions themselves must be processed prior to being usable by higher cognition. Namely, perceptions must be factorized such that they become manipulable. This may include the abstraction of key state features, variables, and structures, and in general represents the process of turning a perception into a belief about the agent's current state. This abstraction process will be informed by long-term memory as well as the agent's current list of goals, tasks, and priorities.

Once perception is factorized, any traditional approach may be used to handle higher order cognitive processes, as translation into symbols is one possible form of factorization. It is also possible to use large language models (LLMs) for some aspects of planning and task scheduling. However, in this case the functions currently handled by a human user in the LLM loop – such as goal setting – would be handled by the metacognitive module in our framework. It is worth noting that deliberative reasoning in humans relies on language. Language acts as a means of abstracting and tokenizing complex ideas, allowing them to be more easily manipulated by the mind. Therefore, it may be reasonable to use transformer-based models as an interface to translate between perceptions and symbols.

As outlined in previous sections, higher order cognitive functions use a prediction-error feedback loop in our framework. The perception cycle is informed by predictions made at this higher level; likewise, low-level predictions are passed to the higher order functions. The metacognitive module, along with attention mechanisms, mediate the degree to which one or the other type of cognition dominates in a given situation – however, both are always active and act in concert. Future work will further explore the interrelationships between different orders of cognition.

The Role of Generative Models

Key to our work is the idea that error (or surprise more generally) is the main trigger to call upon more expensive cognitive resources like deliberation. In our framework, generative models perform a role similar to that of implicit memory: they provide a compressed world model that can be used to generate lossy, reconstructed memories when properly cued. This is in contrast to what we term metalevel memories, which are stable and can be recalled without a contextual prompt. One great advantage of generative models to our use case is their ability to generate predictions – the same mutable quality that makes them poor when precision is needed allows them to be used not only to recollect the past, but to simulate the future without expensive deliberation. This is an essential skill for an intelligent agent, not only in cases where predictions are accurate but also in cases where they fail. When predictions are inconsistent with an agent's observations and the discrepancy cannot be resolved, this is a useful signal that the environment has shifted or the agent's world model was incomplete or inaccurate. In such cases, metalevel control is called upon to determine the best course of action and to engage in additional learning if necessary, and if permitted by the agent's resource bounds.

Discussion

In this section, we summarize some of the research questions that need to be addressed to implement our proposed framework.

1. What are the criteria for subsets of procedural and declarative knowledge to become metalevel knowledge?
2. Does metalevel knowledge capture both abstract and detailed knowledge?
3. How is the metalevel knowledge represented?
4. What is the process for metalevel control to determine if there is a contradiction/conflict between working memory and metalevel memory? How frequently do these checks occur? Why?
5. What are the resource constraints on metaknowledge acquisition and metalevel control? How does this affect performance?
6. When a contradiction or conflict occurs, how does metalevel control determine the the best course of action?
7. Must the metalevel control policy be interpretable by humans?
8. How is the generated data incorporated into the working memory?
9. How is the appropriateness of the metalevel identification of conflict and its response measured in terms of performance improvement and handling non-stationarity?
10. In a multiagent environment, how coordinated do the agents need to be and how does this affect metalevel coordination and knowledge gathering?
11. Our framework requires generative models that learn continuously – how can this be achieved?
12. What is the best way to model the relationship between attention, deliberation, and metacognition (i.e. different levels of executive function)?
13. What is the best way to model the interface between higher order and lower order cognitive functions?
14. Which aspects of higher order cognition should be handled using generative models?

Acknowledgements

The research reported was partially supported by the National Library Of Medicine of the National Institutes of

Health under Award Number: R01LM013327 and a PSC-CUNY 2023 grant. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. It is also supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-2020-0-01789), and the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2023-RS-2023-00254592) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation).

References

A., C. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci*, 36(3): 181–204.

Alexander, G.; Raja, A.; Durfee, E. H.; and Musliner, D. J. 2007. Design Paradigms for Meta-Control in Multi-Agent Systems. In *Proceedings of the AAMAS 2007 Workshop on Metareasoning in Agent-based Systems*, 92–103. Honolulu, HI.

Anderson, J. R.; Bothell, D.; Byrne, M. D.; Douglass, S.; Lebiere, C.; and Qin, Y. 2004. An integrated theory of the mind. *Psychological review*, 111 4: 1036–60.

Bereiter, C.; and Scardamalia, M. 2018. Intentional Learning as a Goal of Instruction. *Knowing, Learning, and Instruction*.

Booch, G.; Fabiano, F.; Horesh, L.; Kate, K.; Lenchner, J.; Linck, N.; Loreggia, A.; Murugesan, K.; Mattei, N.; Rossi, F.; and Srivastava, B. 2021. Thinking fast and slow in AI. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 5042–15046.

Chae, J.; and Kim, J. 2022. Uncertainty-based Visual Question Answering: Estimating Semantic Inconsistency between Image and Knowledge Base. *2022 International Joint Conference on Neural Networks (IJCNN)*, 1–9.

Cheng, S.; Raja, A.; and Lesser, V. 2013. Using Conflict Resolution to Inform Decentralized Learning. In Ito, J.; and Gini, S., eds., *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*, 893–900. St. Paul, MN: IFAAMAS.

Cox, M.; and Raja, A., eds. 2011. *Metareasoning: Thinking about Thinking*. Cambridge, MA.: MIT Press.

Kahneman, D. 2011. *Thinking, fast and slow*. New York: Farrar, Straus and Giroux. ISBN 9780374275631 0374275637.

Kieras, D. E.; and Meyer, D. E. 1997. An Overview of the EPIC Architecture for Cognition and Performance with Application to Human-Computer Interaction. *Hum.-Comput. Interact.*, 12(4): 391–438.

Kim, J.; Myers, K.; Gervasio, M.; and Gil, Y. 2011. Goal-Directed Metacontrol for Integrated Procedure Learning. In *Metareasoning: Thinking about Thinking*. The MIT Press. ISBN 9780262014809.

Laird, J. E.; Lebiere, C.; and Rosenbloom, P. S. 2017. A Standard Model of the Mind: Toward a Common Computational Framework across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics. *AI Magazine*, 38(4): 13–26.

Laird, J. E.; Newell, A.; and Rosenbloom, P. S. 1987. SOAR: An architecture for general intelligence. *Artificial Intelligence*, 33(1): 1–64.

Liu, J.; Shen, Z.; He, Y.; Zhang, X.; Xu, R.; Yu, H.; and Cui, P. 2023. Towards Out-Of-Distribution Generalization: A Survey. arXiv:2108.13624.

Raja, A.; and Lesser, V. 2007. A Framework for Meta-level Control in Multi-Agent Systems. *Autonomous Agents and Multi-Agent Systems*, 15(2): 147–196.