CogPonder: Towards a Computational Framework of General Cognitive Control

Morteza Ansarinia (morteza.ansarinia@uni.lu)
Department of Behavioural and Cognitive Sciences, University of Luxembourg
4366 Belval, Luxembourg

Pedro Cardoso-Leite (pedro.cardosoleite@uni.lu)
Department of Behavioural and Cognitive Sciences, University of Luxembourg
4366 Belval, Luxembourg

Abstract

Current computational models of cognitive control exhibit notable limitations. In machine learning, artificial agents are now capable of performing complex tasks but often ignore critical constraints such as resource limitations and how long it takes for the agent to make decisions and act. Conversely, cognitive control models in psychology are limited in their ability to tackle complex tasks (e.g., play video games) or generalize across a battery of simple cognitive tests. Here we introduce CogPonder, a flexible, differentiable, cognitive control framework that is inspired by the Test-Operate-Test-Exit (TOTE) architecture in psychology and the PonderNet framework in machine learning. CogPonder functionally decouples the act of control from the controlled processes by introducing a controller that acts as a wrapper around any end-to-end deep learning model and decides when to terminate processing and output a response, thus producing both a response and response time. Our experiments show that CogPonder effectively learns from data to generate behavior that closely resembles human responses and response times in two classic cognitive tasks. This work demonstrates the value of this new computational framework and offers promising new research prospects for both psychological and computer sciences.

Keywords: Cognitive Control; Deep Learning; PonderNet; Test-Operate-Test-Exit (TOTE); Cognitive Tests

Cognitive control is a complex construct whose meaning lacks consensus in the literature (Ansarinia, Schrater, & Cardoso-Leite, 2022). One of its key properties is that it allows the cognitive system to regulate its processing to achieve particular outcomes (e.g., inhibit a prepotent response, maintain attentional focus), and this regulation typically has a measurable impact on response times (i.e., control is effortful and takes time). Accordingly, the scientific study of cognitive control has largely focused on how long it takes people to perform tasks (e.g., press a key in response to a light) and on what factors impact those response latencies (e.g., intensity of the light). Developing computational models of cognitive control that replicate human response times remains however a significant challenge in cognitive science.

Various models of response time have been developed in the past, such as the drift diffusion model (DDM; Ratcliff, Smith, Brown, & McKoon, 2016-04). While DDM has its merit, it is limited to certain types of tasks (i.e., binary decision making) and generates data resembling human data distribution rather than being able to directly perform tasks. Meanwhile in machine learning, computational agents can perform complex tasks while overcoming difficult challenges, such as interpreting computations and, crucially, performing human-like control to adapt computation to the task complexity or available resources.

This study aims to develop a computational cognitive control framework that addresses limitations in psychology and machine learning. The desiderata for our framework include agency (being able to perform the task at hand), completeness (accounting for all measured behavior, including response times), versatility (performing a wide range of tasks under a common framework), modularity (flexibility and interpretability of the architecture), and learnability (integration into deep learning frameworks).

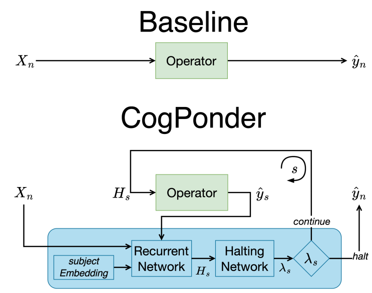

Figure 1: The general CogPonder model template, adaptable for various instances. Our study used a specific implementation for Stroop and N-back tasks with a simple Operator (single dense layer and ReLU activation) and a Controller containing two separate networks: a recurrent (GRUCell) and a halting network. The recurrent network iteratively computes inputs to the Operator, while the halting network estimates the probability of halting at each time step (λ_s, which parameterizes a Bernoulli sample that determines halting the trial at each time step). Though the Operator is slightly task-dependent due to the distinct task requirements (e.g., two versus three responses alternatives), the architecture is extendable for identical configurations across tasks.

CogPonder architecture

We introduce CogPonder, a computational cognitive control framework that fulfills the outlined desiderata. CogPonder is inspired by two primary sources: the PonderNet framework from machine learning and the Test-Operate-Test-Exit (TOTE) framework from psychology. PonderNet is a recent algorithm that adjusts the computational complexity of a neural network based on the complexity of the task and input, allowing the network to use fewer computational steps for simpler tasks (Banino, Balaguer, & Blundell, 2021). It shares similarities with the cognitive model TOTE (Miller, Galanter, & Pribram, 1960), in which computations unfold in cycles with tests evaluating specific conditions and determining whether to halt or continue the process.

CogPonder builds upon an end-to-end off-the-shelf model called the Operator, which simply takes an input and outputs a response. The key design principle behind CogPonder is to wrap the Operator within a local virtual environment governed by the Controller, which intercepts the inputs and outputs, determining what inputs are fed to the Operator and what output is emitted at a given time by the system. The Controller, implemented similarly to PonderNet, adjusts the system's computational complexity and determines its response time by computing the probability of halting at each time step within a given trial. Unlike comparable deep learning architectures for human response times which relies on decision uncertainty (e.g., RTNet; Rafiei & Rahnev, 2022), CogPonder's response time is determined by computational requirements. Furthermore, CogPonder aims to align computation time with human behavior, emphasizing TOTE's building blocks metaphor and furthering our understanding of cognitive control theory.

Evaluation

We evaluated a CogPonder instance using human data by aiming to align the timing of the model's output with human response times. As a first test, we examine a single CogPonder agent independently performing two cognitive control tasks. This proof of concept is vital, as it shows how tasks previously studied in isolation can now be investigated within a common computational framework.

Method We train a CogPonder agent to compute responses at each time step and output final choices, thus generating for this agent data that has the same structure as human data (i.e., trial-level data; Defossez et al., 2020), which enables the direct comparison of agent and human behavior. The agent aligns with human behavior by minimizing the loss function, L_{total}, which comprises two terms that are weighted by the hyperparameter β: $L_{\text{total}} = L_{\text{response}} + \beta L_{\text{time}}$. The first term, L_{response}, called *reconstruction loss*, aims to match agent choices to human choices using the cross entropy loss function. The second term, L_{time}, called *regularization loss*, aims to align the agent's number of computational steps within each trial with human response times using KL divergence. It is important to note that our method contrasts agent's halting steps with participant response times. More specifically, we track the agents' number computational steps (n) rather than elapsed computation time (i.e., seconds), as the latter varies depending on the hardware. To compute L_{time}, we convert response times from seconds to steps using a step duration hyperparameter.

Data We evaluated CogPonder using the Self-Regulation Ontology dataset, containing behavioral data from 521 participants who completed cognitive tests (for further details on the datasets, see Eisenberg et al., 2019). Specifically, we focused on participants who completed the Stroop and 2-back tests, chosen for their relevance to cognitive control despite their differences in stimuli, task instructions, cognitive processes, and response options. The Stroop task required participants to identify the ink color of words while ignoring the words themselves which are color names that are congruent or incongruent with their ink. The 2-back task involved reporting if a letter matched the one presented two letters earlier. For both tasks, trial data included stimulus, response, and response time. The data, which represented a time series of trials, was split into 75% training and 25% test sets.

Experimental setup Model training involved up to 10,000 epochs using the Adam optimizer (learning rate of 0.01). All parameters were tuned automatically, except for the step duration hyperparameter, set manually at 20 milliseconds (future iterations will estimate this using a dedicated validation set). Furthermore, as customary (Luce, 1986), a non-decision time hyperparameter was incorporated into the model, accounting for duration factors unrelated to decision-making.

Results The first goal of this study was to assess the extent to which a CogPonder model behaves like a human after being trained with human data. We compared the average accuracy and response time of human and CogPonder agents in Stroop and 2-back tests (Figure 2). Our results show that CogPonder is able to capture broad patterns in human data, producing similar accuracy and response times across trial types.

We then assessed CogPonder's ability to reproduce more nuanced human phenomena by analyzing 1) average accuracy and response time as a function of experimental conditions (e.g., congruent versus incongruent Stroop trials) and 2) response time distributions in the two tests. Although there are some differences, CogPonder approximately mimics human's average accuracy and response time (Figure 2, first and second rows), as well as similar response time distributions (Figure 2, third row).

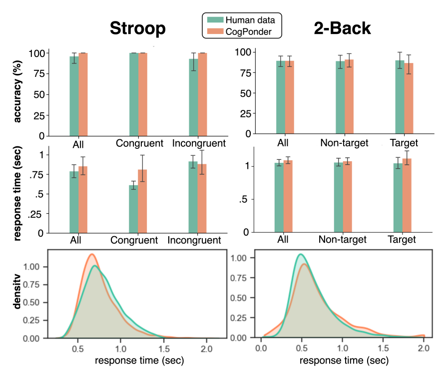

Figure 2: Comparing CogPonder agent and human agent

Concluding remarks

CogPonder is an initial effort to create a general computational cognitive control framework suitable for various use-cases, in particular modeling behavior across cognitive test batteries. CogPonder's performance demonstrates its ability to align with human behavior, making it valuable for studying cognition across multiple cognitive tests — this framework can easily be extended to accommodate a wider range of tasks. Importantly, CogPonder breaks the "complexity of behavior ceiling" relative to existing approaches, making it potentially applicable not only to multiple simple tests but also to complex ones.

Future work will extend the CogPonder implementation and evaluation procedures, test additional cognitive tasks, patterns of inputs and outputs (see Defossez et al., 2020), learn computation hierarchies of tasks following the TOTE's building block metaphor, and further investigate CogPonder's ability to generalize across tasks.

Acknowledgments

This research was supported by the Luxembourg National Research Fund (ATTRACT/2016/ID/11242114/DIGILEARN). **MA** wrote the initial draft and developed the first prototype. **PCL** edited the initial draft, supervised the project, and edited/reviewed the manuscript.

References

Ansarinia, M., Schrater, P., & Cardoso-Leite, P. (2022). *Linking theories and methods in cognitive sciences via joint embedding of the scientific literature: The example of cognitive control.*

Banino, A., Balaguer, J., & Blundell, C. (2021). PonderNet: Learning to Ponder. doi: 10.48550/ARXIV.2107.05407

Defossez, A., Ansarinia, M., Clocher, B., Schmück, E., Schrater, P., & Cardoso-Leite, P. (2020). The structure of behavioral data. doi: 10.48550/ARXIV.2012.12583

Eisenberg, I. W., Bissett, P. G., Zeynep Enkavi, A., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. , *10*(1), 2319. doi: 10.1038/s41467-019-10301-1

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization* (No. no. 8). Oxford University Press ; Clarendon Press.

Miller, G., Galanter, E., & Pribram, K. (1960). *Plans and the structure of behavior.*

Rafiei, F., & Rahnev, D. (2022). *RTNet: A neural network that exhibits the signatures of human perceptual decision making.* doi: 10.1101/2022.08.23.505015

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016-04). Diffusion Decision Model: Current Issues and History. , *20*(4), 260–281. doi: 10.1016/j.tics.2016.01.007