

Towards metacognitive agents: integrating confidence in sequential decision-making

B. Pesquet¹ and F. Alexandre¹

1- INRIA, University of Bordeaux, CNRS, Bordeaux INP - France

Abstract. In natural cognition, confidence is used to evaluate the quality of decisions and adapt one's behavior to the task at hand. For now, artificial agents lack this kind of metacognitive ability and interact with their environment in a purely reactive way. Inspired by recent findings about the cognitive modeling of confidence, we propose a novel architecture for sequential decision-making. It combines an evidence accumulation model with a metacognitive module that computes and exploits confidence to tune the decision process. The model has been assessed on a perceptual decision-making task, showing promises for more flexible artificial agents and a possible path towards artificial metacognition.

1 Introduction

Everyday, we make a myriad of choices regarding many topics, big and small. Decision-making, the cognitive process leading to a decision, is one of the key skills to adapt to one's environment and reach one's goals. At the cognitive level, a number of these decisions are taken in a reactive and implicit way, using knowledge and associations learned through previous experiences. For example, when asked if $2 + 2 = 5$, chances are you would answer immediately without any explicit reasoning. But some other decisions are tougher to make, either because of the imperfections in the information guiding the choice, or because of the stakes involved [1]. Imagine, for example, that you are preparing yourself for an upcoming test. Before the big day, deciding whether or not you understand the material well enough to stop studying is not an easy choice to make. After taking the test, you might reflect (more or less anxiously) on your answers, trying to assess your performance level. Both cases imply a form of self-evaluation regarding your cognitive processes and an explicit deliberation which might inhibit a too simple reactive decision and promote instead a contextually more adapted decision. The ability to reflect on, evaluate and control one's mental functions is called *metacognition*. This practice of "thinking about thinking" is a key skill to adapt to complex problems and changing environments.

Concerning the first aspect of reactive and implicit decision making, Artificial Intelligence (AI) has proposed several approaches, including reinforcement learning with a slow learning based on reward prediction errors. Similarities at both the behavioral and computational levels have been reported with reactive decision making in animals [2]. AI has only recently begun to consider explicit and deliberative decision making, mainly by proposing ways to emulate deliberation, like with Chain-of-Thought prompting in Large Language Models [3]. Yet, this kind of method is still very preliminary, with an unreasonable computational

cost and far from some human characteristics (e.g., flexibility, explainability) [4]. In addition, the stage of self-evaluation playing a major role before the explicit control of cognitive processes is largely neglected in modern AI.

Another line of research more associated with experimental psychology has proposed to explicitly model the sequential deliberative process of decision making as accumulation of evidence [5]. Initially designed for binary choices and not considering the learning phase, this model family has been extended to ease the implementation of coupled learning and decision in cognitive agents [6] and studies to assess its biological plausibility have been proposed [7]. Such models are very interesting because they can provide quantitative explanations of psychometrics and chronometrics of decision making, but under the condition that an accepted error rate can be given a priori to the model.

Artificial agents lack the capacity to reflect upon themselves, to question their choices and behaviors. Augmenting them with metacognitive abilities could be essential to improve their performance, but also their acceptability. In particular, the ability to evaluate one's ability to make a decision, that is also called confidence, is a necessary step. A simple but efficient way of implementing such a mechanism in a sequential decision process is described below and evaluated on a simple perceptual task.

2 Related work

2.1 Sequential decision-making

A decision is a deliberative process that results in the commitment to a categorical proposition [8]. Decision making is generally a *sequential* process, either because we need time to collect informative cues, one after the other (like a detective collecting pieces of evidence) or because we need time to process each cue, one after the other (like a chess player).

A popular class of models assumes that a decision maker accumulates pieces of evidence for each alternative until a threshold is reached for one alternative. Such accumulation-to-threshold models are known as *sequential sampling models* or *evidence accumulation models* (EAM). One of the first and most popular EAMs is the Diffusion Decision Model (DDM), also called Drift-Diffusion Model, originally designed in the 1970's [9] to study binary choices.

This class of models is used to study the time constraints of decision by accumulation [10]. The intuitively understandable balance between response time and accuracy is called the *speed/accuracy trade-off* (SAT) [11].

In their simplest form, evidence accumulation models have three parameters: the boundary separation a that implements the SAT, the starting point z that represents the pre-decision bias in favor of one of the responses, and the drift rate v that models the speed of evidence accumulation processing.

Many subsequent works have refined this seminal model to account for multiple choices [12], change of mind [13] or impacts of learning on the decision outcome [14], [15].

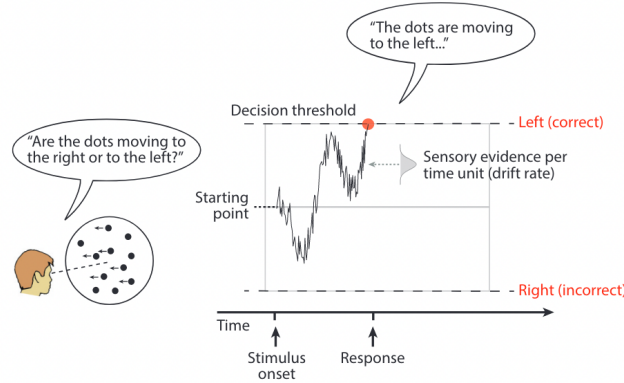


Fig. 1: Illustration of the Diffusion Decision Model [10]

2.2 Confidence in decision-making

Everyone knows intuitively what confidence is about, yet it is seldom defined explicitly. In the broadest sense, confidence quantifies a degree of belief in something or someone [16], where a belief can be defined as a feeling of certainty about a proposition (i.e. a statement or a decision). It is a subjective, conscious experience. Confidence is fundamentally linked to its object: a thought, a choice, an external actor, etc.

In decision-making, confidence can be seen as the subjective estimate of decision quality [17]. More formally, it can be defined as the probability that a choice is correct given the evidence [18]. Confidence is a form of certainty. A key difference is that contrary to confidence, (un)certainties are decision independent [1]. Confidence quantifies the degree of certainty associated with a decision.

When using an evidence accumulation model to study the decision process, a simple and natural way of evaluating confidence is by computing the difference between accumulators for each alternative. The larger the gap, the higher the confidence should be, reflecting a more robust decision. Furthermore, this approach is biologically grounded [19].

3 Proposed model

Taking into account the previous literature on confidence and human decision-making, we designed a model for an artificial agent including two components:

- A *decision module* which uses an evidence accumulation model to select the chosen alternative.
- A *metacognitive module* which tunes the decision process through adjusting the decision hyperparameters (threshold and rate of evidence integration; the starting point is not considered in the task here).

The decision module’s evidence accumulation is based on a *race model*. Evidence favoring each alternative is accumulated separately, and the decision is taken once the first accumulator reaches the threshold.

The metacognitive module uses the difference between accumulators to compute the confidence associated with the decision. Confidence is then used to tune the decision process, implementing the metacognitive ability of switching between different strategies. For example, detecting a low confidence value should trigger a more cautious approach and consequently a longer time of evidence accumulation; conversely, being more confident should lead to a quicker decision.

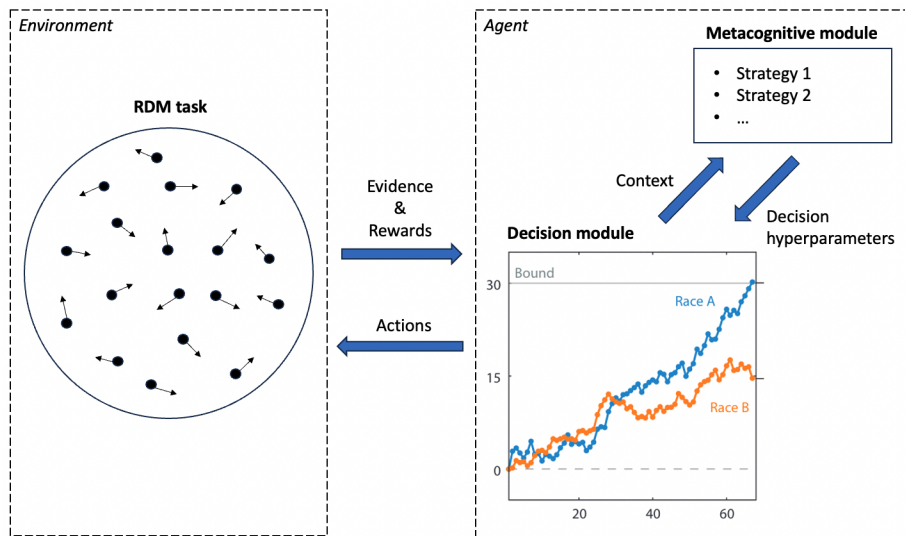


Fig. 2: Task and agent model

The model was trained on an adaptation of the Random Dot Motion (RDM) task. This well-known psychophysical experiment continuously presents a set of moving dots to the subject, who has to make a decision regarding their movement: do the majority of dots move to the left or to the right? The experimenter can modulate the difficulty of the task through several factors, notably the coherence of the dot cloud (percentage of dots moving in the same direction).

4 Experimental validation

The model was implemented in Python and trained on a custom Gymnasium environment mimicking the RDM task. Two variations of the task were used: one with a strong coherence between dots and another with more random movements.

At each timestep of a training episode, information about dots is provided

to the model as raw pixel values. Integrating this evidence in the corresponding accumulator, the agent can decide to wait for more data or make a decision regarding the dot movement (which ends the episode). After each episode, rewards are considered a posteriori by the model to adjust the metacognitive process.

The first results demonstrate how the metacognitive module alters the decision process. They indicate that the agent model is able to implement a Speed-Accuracy Tradeoff [11], a hallmark of human cognition.

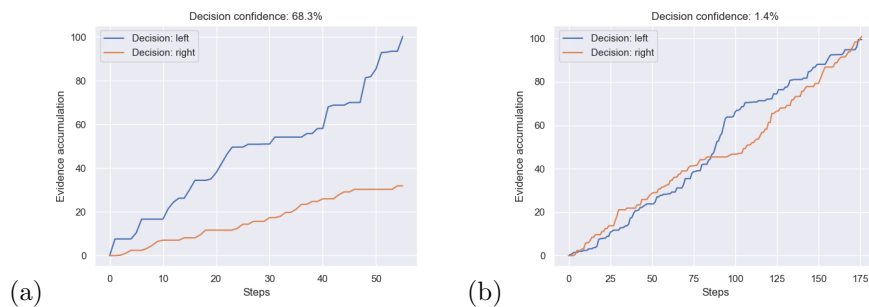


Fig. 3: Experimental results. (a) Easy environment: high confidence and short decision process. (b) Harder environment: low confidence and longer process.

5 Conclusion and perspectives

Human decision-making is of sequential nature and characterized by our ability to reflect upon ourselves and adapt our behavior accordingly. A growing body of work at the crossroads of psychology and computational neuroscience studies the role of confidence in the decision process, and its formation in the brain.

Drawing from these premises, we defined a model for a metacognitive artificial agent able to evaluate its confidence and use it to adapt its decision process on-the-fly. We tested this model on a classic perceptual decision task. First results show the positive (and expected) impact of metacognition on the agent's behavior.

Ongoing works are carried out to consolidate the first results regarding confidence evaluation and use in decision making. Further developments will be needed to refine and extend the cognitive architecture to more general aspects of metacognition, towards autonomous and flexible behavior of an artificial agent.

References

- [1] Stephen M. Fleming. Metacognition and Confidence: A Review and Synthesis. *Annual Review of Psychology*, 75(Volume 75, 2024):241–268, January 2024. Publisher: Annual Reviews.
- [2] R. E. Suri and Wolfram Schultz. A Neural Network Model with Dopamine-Like Reinforcement Signal that Learns a Spatial Delayed Response Task. *Neuroscience*, 91(3):871–890, July 1999.

- [3] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed H. Chi, Quoc V. Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22, pages 24824–24837, Red Hook, NY, USA, April 2024. Curran Associates Inc.
- [4] Brenden M. Lake, Tomer D. Ullman, Joshua B. Tenenbaum, and Samuel J. Gershman. Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40, 2017. Publisher: Cambridge University Press.
- [5] Roger Ratcliff, Philip L. Smith, Scott D. Brown, and Gail McKoon. Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, 20(4):260–281, April 2016.
- [6] Don van Ravenzwaaij, Scott D. Brown, A. a. J. Marley, and Andrew Heathcote. Accumulating advantages: A new conceptualization of rapid multiple choice. *Psychological Review*, 127(2):186–215, March 2020.
- [7] Bhargav Teja Nallapu and Frédéric Alexandre. Advantage Race Diffusion mechanism to study the dissociated roles of Orbitofrontal Cortex subregions in instrumental learning. In *The 5th quadrennial meeting on orbitofrontal cortex function (OFC 2024)*, April 2024.
- [8] Joshua I. Gold and Michael N. Shadlen. The Neural Basis of Decision Making. *Annual Review of Neuroscience*, 30(1):535–574, July 2007.
- [9] Roger Ratcliff. A theory of memory retrieval. *Psychological Review*, 85:59–108, 1978.
- [10] B.U. Forstmann, R. Ratcliff, and E.-J. Wagenmakers. Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions. *Annual Review of Psychology*, 67(1):641–666, January 2016.
- [11] Richard P. Heitz. The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience*, 8, June 2014.
- [12] Don Van Ravenzwaaij, Scott D. Brown, A. A. J. Marley, and Andrew Heathcote. Accumulating advantages: A new conceptualization of rapid multiple choice. *Psychological Review*, 127(2):186–215, March 2020.
- [13] Arbora Resulaj, Roozbeh Kiani, Daniel M. Wolpert, and Michael N. Shadlen. Changes of mind in decision-making. *Nature*, 461(7261):263–266, September 2009.
- [14] Laura Fontanesi, Sebastian Gluth, Mikhail S. Spektor, and Jörg Rieskamp. A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review*, 26(4):1099–1121, August 2019.
- [15] Steven Miletić, Russell J Boag, Anne C Trutti, Niek Stevenson, Birte U Forstmann, and Andrew Heathcote. A new model of decision processing in instrumental learning tasks. *eLife*, 10:e63055, January 2021.
- [16] Florent Meyniel, Mariano Sigman, and Zachary F. Mainen. Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron*, 88(1):78–92, October 2015.
- [17] Jeroen Brus, Helena Aebersold, Marcus Grueschow, and Rafael Polania. Sources of confidence in value-based choice. *Nature Communications*, 12(1):7337, December 2021.
- [18] Alexandre Pouget, Jan Drugowitsch, and Adam Kepecs. Confidence and certainty: Distinct probabilistic quantities for different goals. *Nature Neuroscience*, 19(3):366–374, March 2016.
- [19] Adam Kepecs, Naoshige Uchida, Hatim A. Zariwala, and Zachary F. Mainen. Neural correlates, computation and behavioural impact of decision confidence. *Nature*, 455(7210):227–231, September 2008.