Reducing Intuitive-Physics Prediction Error through Playing

Olivier L. Georgeon[1,2][0000−0003−4883−8702] and Paul Robertson[3][0000−0002−4477−0379]

[1] UR CONFLUENCE: Sciences et Humanites (EA 1598), UCLy, France
ogeorgeon@univ-catholyon.fr
[2] SyCoSMA, LIRIS, CNRS, Villeurbanne, France
[3] DOLL Labs, Lexington, MA, USA
paulr@dollabs.com

Abstract. We present a method for an autonomous robot to generate behaviors to calibrate its intuitive-physics engine also known as the "Game Engine in the Head" (GEITH). At the beginning of each interaction cycle, the robot uses its GEITH to run a simulation to compute predicted sensory signals. For each sensor, prediction error is the difference of the predicted sensory signal minus the actual sensory signal received at the end of the interaction cycle. Results show that over a few tens of interaction cycles, the robot improves its GEITH calibration and decreases its prediction errors. Moreover, the robot generates behaviors that human observers describe as playful.

Keywords: Active infernce · constructivist learning · enaction · intrinsic motivation · robotics · core knowledge.

1 Introduction

It is widely believed that cognitive beings possess some kind of a *world model* that they use to generate intelligent behaviors. How they construct, maintain, and use this world model remains, however, an open question in cognitive science and artificial intelligence.

The Partially Observable Markov Decision Process (POMDP) literature has proposed a broad range of methods to infer a *belief state* in a partially observable process. The belief state amounts to the agent's world model of the environment that the agent can only partially observe. If the state transition function and the observation function are known *a priori*, the problem of computing the belief state has been mathematically solved [1]. It was, however, also proven that the implementation of the solution becomes intractable as the set of states and observation grows.

Karl Friston and his research group have proposed Active Inference [11, e.g.] as a method to infer the world model by minimizing *free energy* [4]. The world model is represented as a probability distribution that gives the probability of being in any particular state among all the possible states of the world. This method, however, requires that the set of all worlds states and the relations between states and observations be known *a prior*. The computational requirements to compute the free energy, continue to pose significant challenges to implementing autonomous robots capable of learning a model of the open world.

The POMDP and active inference literature led us to hypothesize that inferring the world model through experience of interaction requires prior assumptions to reduce complexity [7]. The present study examines how the "Game Engine In The head" (GEITH) can work as a suitable prior assumption that an autonomous robot could use to reduce its prediction errors.

Joshua Tenenbaum and his research group have proposed the GEITH [2] as the capacity of cognitive beings to simulate basic dynamics of physics and interactions. In mammals, the GEITH would rest upon brain structures that are partially predefined by genes and then completed through ontogenetic development. Similarly, it is possible to endow artificial agents and robots with a predefined software game engine, and expect them to refine the parameters of their game engine as they test their predictions in the world.

This paper proposes an approach to designing autonomous agent that can refine their GEITH through their lifetime. The refinement of the game engine is measured through two methods. The first is performed by the robot itself by measuring the prediction error of sensory signals. Decrease in prediction errors shows improvement of the game engine. The second is performed by the experimenter by assessing whether the game engine parameters converge towards a target range that indicates that the robot managed to calibrate its GEITH.

2 Our hypothesis

We comply with active inference theory in several regards. Firstly, we do not consider sensory signals to be *representational* of the environment's state. Different sensory signals may come out of the same environment state depending on the agent's action. This implies a "conceptual inversion" of the interaction cycle in which action comes first and sensory signal comes second as an *outcome* of action. Secondly, we do not provide the agent with presupposed ontological knowledge about entities in the environment. The agent must infer the existence and properties of entities by itself through patterns of interactive experience. This view can be tracked back to Whitehead's process philosophy in which phenomenal experience involves abstracting entities out of events [14]. Thirdly, no extrinsic goal is encoded in the agent in the form of goal states that the agent should search based on reward or other criteria. We nonetheless accept to encode some *prior preference distribution* over tuples $\langle$action, outcome$\rangle$ that encode "innate" preferences of interactions. In short, the agent has no *rewarding world states* but has *rewarding interactions* (positive or negative). For a deeper examination of these principles in relation with the active inference literature, we refer the reader to our previous article [7].

We also adopt *prediction error* as a measure of the quality of the agent's world model. Indeed, to survive, cognitive beings should avoid bad predictions as much as they can. For reasons introduced in Section 1, however, we are not using gradient descent of prediction error (or surprise) as a motivational principle to drive the learning process. As we will develop, our agent is not always driven by a value optimization process; it may sometimes enact selfless behaviors. For us, prediction-error reduction is not a means to improve the world model but a consequence of its improvement.

We are using a cognitive architecture that we designed previously based on sensorimotor and enactive principles [7]. The present article reports the integration of the new GEITH module within this cognitive architecture as illustrated in Figure 1. The GEITH supports the simulation of interactive behaviors before their selection by the cognitive architecture and their enaction by the robot. At the beginning of each interaction cycle, the simulation computes the *predicted outcome*. At the end of the interaction cycle, the predicted outcome is compared with the *actual outcome* to compute the prediction error. We investigate the core elements of the GEITH that are needed for the agent to reduce prediction error.

We draw inspiration from studies on *core knowledge* in the brain of animals and human infants. For example, Elizabeth Spelke and her colleagues argued for the existence of two core systems of geometry that "evolved even before the emergence of the human species": "The *core navigation system* captures absolute distance and sense [..] but not relative length or angle; the *core form analysis system* does the reverse" [12, p. 2789]. We start by implementing the minimal requirements she deems necessary for both systems, namely the ability to handle points and lines in spatial memory, the foundational elements of Euclidean geometry.

Our cognitive architecture encodes behaviors as *composite interactions* which are sequences of *primitive interactions*. A primitive interaction is a specification of a *control loop* that involves actuator commands, expected sensory feedback, spatio-temporal attributes, termination conditions, and termination outcome. Section 3 will provide examples. The GEITH may consider some outcomes as resulting from the interaction with "something" in the environment. In this case, the GEITH instantiates a data structure called a *phenomenon*[4] and localizes this phenomenon at the position of the interaction in spatial memory. Next, the GEITH simulates subsequent interactions with phenomena to predict future outcomes. The present study focuses on the simplest possible phenomenon: a point on the two-dimensional floor.

We seeded the cognitive architecture with "innate" composite interactions (Fig. 1, top-center) that cause the robot to explore the environment and to interact with points encountered on the floor. Here we do not examine the learning of new composite interactions (this was the object of other studies [8]) but only the refinement of the GEITH parameters to reduce prediction error. The cognitive architecture uses variables that represent the robot's *emotional state* to selects composite interactions to try to enact.

[4] Common-sense usage of the term *phenomenon*: "something" that a cognitive being perceives in the environment. Technically: "any useful grouping of a subset of spatio-temporal patterns experienced by an agent in an environment" [13, p. 8].

We use Hugo Lövheim's "cube of emotions" [9] as a basic emotional model based on three neurotransmitters: dopamine (DA), serotonin (5-HT), and nor-adrenaline (NA) (Fig. 1, top right). This model associates dopamine with pleasure and reward-seeking behavior, serotonin with well-being and playful behavior, and nor-adrenaline with arousal and stress responses. It has been successfully used for simple emotional robotics. Our robot visually indicates the predominant neurotransmitter level using an intuitive color code developed by Max Talanov and his team: green for dopamine, white for serotonin, red for noradrenaline, and blue when all three neurotransmitter levels are low [3].

The robot must follow a chain of causality to explain the prediction errors [13].

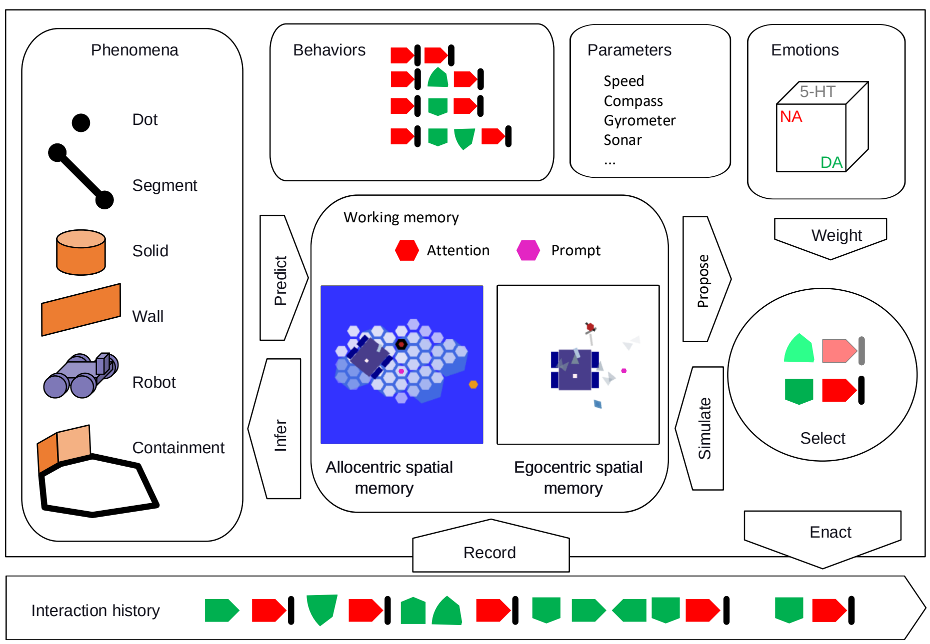

Fig. 1. The game engine within the robot's cognitive architecture. Bottom: the history of interactions enacted over time. Center: the working memory that implements the game engine and proposes future behaviors. Left: the types of phenomena inferred trough interactive experience. Top center: set of behaviors. Top right: The three-dimensional emotional state based on dopamine (DA), serotonin (5-HT), and nor-adrenaline (NA) levels. Right: the decider selects the next behavior based on the emotional state and the expected outcome predicted from simulation.

3 Experiment

We designed a robotic platform called Petitcat based on the "robot car" commercialized by Osoyoo [10]. For this experiment, we added an Inertial Measurement Unit (IMU) and an RGB LED as the emotion indicator (Fig. 2).

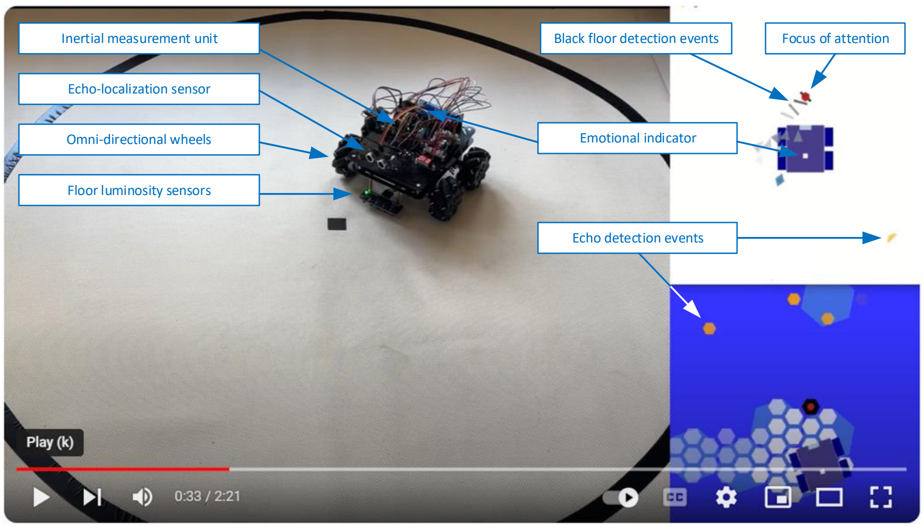

Fig. 2. Screenshot of a video example run [5]. Left: Petitcat playing with a point on the floor. Top right: Petitcat's egocentric memory. Black segments: black point detection events. Bottom right: allocentric memory. Black hexagon: the black point used as a point of reference. Yellow hexagons: echo measured with the sonar. Red hexagon: the robot's focus of attention.

The C++ software running on the Arduino board controls the enaction of primitive interactions. A personal computer implements the GEITH and the cognitive architecture that remote controls the robot through wifi. The cognitive architecture selects the primitive interaction to try to enact and sends it to the robot. The robot enacts it and sends the outcome back to the PC. The code is open source and shared online [6].

For example, the primitive interaction `move forward` consists of activating the four wheels at equal speed. The control loop checks luminosity of the floor, the acceleration, and the distance of echo signals. The spatio-temporal attribute is the duration or the distance to travel. The termination conditions are reaching the duration or detecting an obstacle that interrupts the traveling. The termination outcome is the actual duration of traveling and the possible detection of obstacle.

The levels of all three neurotransmitters can vary from 0 to 100 and are initialized at 50. DA prevails when DA = 5-HT = NA. Prevalence of DA makes

Petitcat initially select the `move forward` behavior that is intrinsically rewarding. When he encounters a new object, 5-HT increases to MAX-5-HT. Prevalence of 5-HT triggers the selection of playful interactions with this object. On each `play` interaction, if the prediction errors do not decrease (i.e., prediction does not improve), 5-HT decreases by 1. When 5-HT decreases below or equal to DA, exploration behaviors are again selected meaning that Petitcat disinterests from the object and goes exploring new destinations. During play, when Petitcat unexpectedly fails to detect the object, NA increases to MAX-NA, which causes the selection of `search` behaviors. DA then decreases of 1 on each interaction or is reset to MIN-NA when Petitcat finds the object again.

4 Results

Several videos are available online. Here we analyze a representative run recorded on video [5].

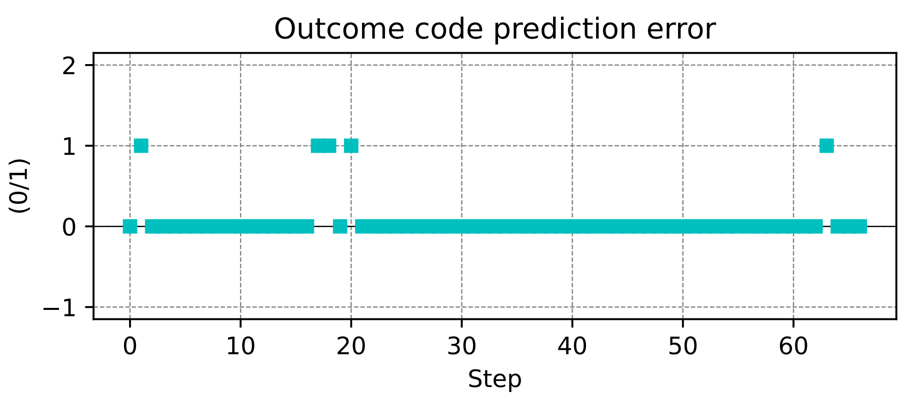

Fig. 3. Prediction error of black phenomenon detection. Step 2: the robot did not expect to detect the point. Steps 17: the robot expected to detect the point while translating forward but missed it. Steps 18: the robot expected to not detect the point while turning around but detected it. Step 20: the robot did not predict detecting the point through simulation but it did. Step 65: The robot did not expect to detect the surrounding arena.

5 Conclusion

We are not claiming the robot can actually *experience* emotions let alone have sentience. His internal model of emotions, nonetheless, helps generate behaviors that human observers easily interpret as lifelike. The emotional indicator facilitates this interpretation.

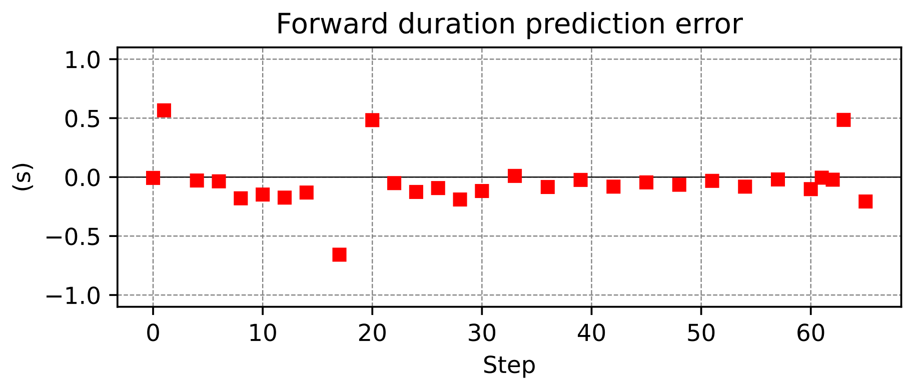

Fig. 4. Move forward duration prediction error. Step 2 and 20: the forward translation was unexpectedly interrupted by the point detection. Step 17: the forward duration was longer than expected because the robot did not detect the point. From Step 21 to 64: forward duration prediction error slightly decreases. Step 65: the forward translation was unexpectedly interrupted by the detection of the arena border.

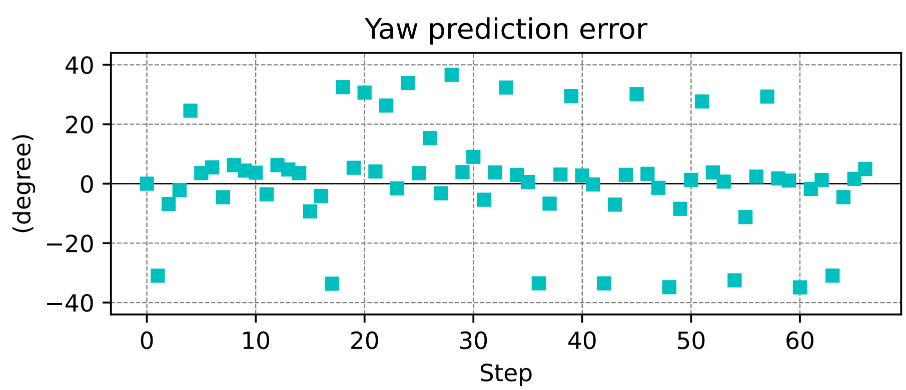

Fig. 5. The yaw prediction error shows no significant trend when we do not distinguish between the different kinds of interactions.

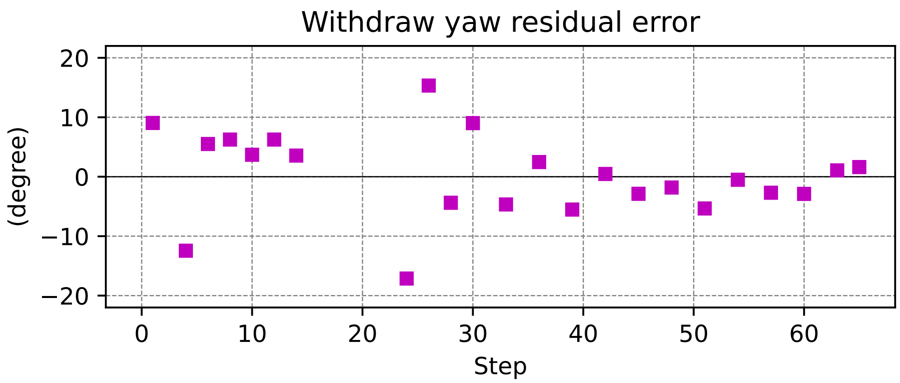

Fig. 6. The yaw residual error during interactions in which the robot detects the point decreases significantly after Step 26.

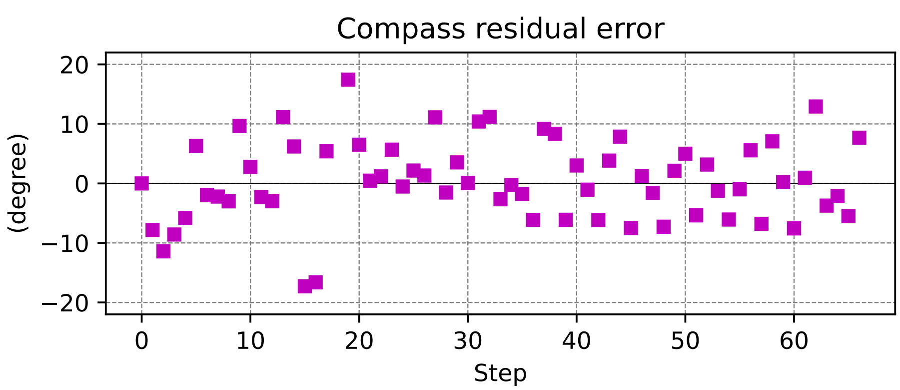

Fig. 7. The compass residual error decreases after step 20 when the robot starts circling around the point. It nonetheless remains noisy due to sensor imprecision. The sliding average over 10 interactions tends to 0.8° and the standard deviation to 7.4°.

References

1. Åström, K.J.: Optimal control of markov processes with incomplete state information. Journal of mathematical analysis and applications **10**(1), 174–205 (1965)
2. Battaglia, P.W., Hamrick, J.B., Tenenbaum, J.B.: Simulation as an engine of physical scene understanding. Proceedings of the National Academy of Sciences **110**(45), 18327–18332 (2013). https://doi.org/10.1073/pnas.1306572110
3. Chebotareva, E., Safin, R., Shafikov, A., Masaev, D., Shaposhnikov, A., Shayakhmetov, I., Magid, E., Zilberman, N., Gerasimov, Y., Talanov, M.: Emotional social robot "emotico". In: 2019 12th International Conference on Developments in eSystems Engineering (DeSE). pp. 247–252. IEEE (2019). https://doi.org/10.1109/DeSE.2019.00054
4. Friston, K.: The free-energy principle: a unified brain theory? Nature Reviews Neuroscience **11**(2), 127–138 (2010). https://doi.org/10.1038/nrn2787
5. Georgeon, O.L.: Petitcat calibrates its intuitive physics engine (2024), https://youtu.be/4wF-OeYCcYI
6. Georgeon, O.L.: Petitcat project repository (2024), https://github.com/UCLy/INIT2/
7. Georgeon, O.L., Lurie, D., Robertson, P.: Artificial enactive inference in three-dimensional world **86**, 101234 (2024). https://doi.org/10.1016/j.cogsys.2024.101234
8. Georgeon, O.L., Riegler, A.: CASH only: Constitutive autonomy through motorsensory self-programming. Cognitive Systems Research **58**, 366–374 (Dec 2019). https://doi.org/10.1016/j.cogsys.2019.08.006
9. Lövheim, H.: A new three-dimensional model for emotions and monoamine neurotransmitters **78**(2), 341–348 (2024). https://doi.org/10.1016/j.mehy.2011.11.016
10. Osoyoo: M2.0 metal chassis mecanum wheel robotic (2022), https://osoyoo.com/2022/07/05/v2-metal-chassis-mecanum-wheel-robotic-for-arduino-mega2560-introduction-model-2021006600/
11. Smith, R., Friston, K.J., Whyte, C.J.: A step-by-step tutorial on active inference and its application to empirical data. Journal of Mathematical Psychology **107**, 102632 (2022). https://doi.org/10.1016/j.jmp.2021.102632
12. Spelke, E.S., Lee, S.A.: Core systems of geometry in animal minds **367**(1603), 2784–2793 (2012). https://doi.org/10.1098/rstb.2012.0210
13. Thórisson, K.R.: The 'Explanation Hypothesis' in general self-supervised Learning. International Workshop in Self-Supervised Learning (2021)
14. Whitehead, A.N.: Process Philosophy. Macmillan (1929)