Reducing Prediction Error by Refining the Game Engine In The Head

Olivier L. Georgeon[1[0000−1111−2222−3333]], Second Author[2,3[1111−2222−3333−4444]], and Third Author[3[2222−−3333−4444−5555]]

[1] UR CONFLUENCE: Sciences et Humanites (EA 1598), UCLy, France
`ogeorgeon@univ-catholyon.fr`
[2] Springer Heidelberg, Tiergartenstr. 17, 69121 Heidelberg, Germany
`lncs@springer.com`
http://www.springer.com/gp/computer-science/lncs
[3] ABC Institute, Rupert-Karls-University Heidelberg, Heidelberg, Germany
`{abc,lncs}@uni-heidelberg.de`

Abstract. We present a method for a robot to generate behaviors to refine its "Game Engine in the Head" (GEITH). The behaviors result in posterior sensory signals that carry information to refine the GEITH parameters. 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 hundred interaction cycles, the robot manages to satisfactorily calibrate its GEITH, and prediction errors decrease. Moreover, the robot generates behaviors that human observers describe as playful.

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

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 and maintain this world model remains, however, an open question in cognitive science and artificial intelligence.

Karl Friston and his research group have proposed Active Inference [4, e.g.] as a method to infer the world model by minimizing free energy [3]. The world model is represented as a probability distribution $\mu = P(S)$ over the set S of possible world states. This method iteratively updates the world model after interaction through gradient descent of free energy. The variational free energy amounts to the divergence between two probability distributions: the estimated world model μ and the joint probability distribution of observations and world states called the *generative model*. This method, however, requires that the set of states S and the relations between states and observations be known *a prior*.

Moreover, the high computational needs to compute the free energy and the high number of interaction cycles to converge to a useful world model makes this method inapplicable in our case of a robot interacting with the open world.

The Partially Observable Markov Decision Process (POMDP) literature proposes 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 also proven that the implementation of the solution becomes intractable as the set of states and observation grows. In the absence of these presupposition, the problem of inferring belief states in POMDPs does not lend itself to a mathematical analysis.

The active inference and the POMDP literature suggests that inferring the world model through experience of interaction requires prior assumptions to reduce complexity. The present study proposes the hypothesis that the "Game Engine In The head" (GEITH) can work as a suitable prior assumption.

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.

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 show 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.

fqfdqsfdqs fdsqfdsq fdqs

fdqsfdqs

fdqs

fqsfqs

fdqsfdqs

fdqsfqs

2 Experimental setup

2.1 The robotics platform

2.2 The Game Engine In The Head

3 Behavior generation

4 Results

5 Conclusion

6 First Section

6.1 A Subsection Sample

Please note that the first paragraph of a section or subsection is not indented. The first paragraph that follows a table, figure, equation etc. does not need an indent, either.

Subsequent paragraphs, however, are indented.

Sample Heading (Third Level) Only two levels of headings should be numbered. Lower level headings remain unnumbered; they are formatted as run-in headings.

Sample Heading (Fourth Level) The contribution should contain no more than four levels of headings. Table 1 gives a summary of all heading levels.

Table 1. Table captions should be placed above the tables.

Heading level	Example	Font size and style
Title (centered)	**Lecture Notes**	14 point, bold
1st-level heading	**1 Introduction**	12 point, bold
2nd-level heading	**2.1 Printing Area**	10 point, bold
3rd-level heading	**Run-in Heading in Bold.** Text follows	10 point, bold
4th-level heading	*Lowest Level Heading.* Text follows	10 point, italic

Displayed equations are centered and set on a separate line.

$$x + y = z \tag{1}$$

Please try to avoid rasterized images for line-art diagrams and schemas. Whenever possible, use vector graphics instead (see Fig. 1).

Theorem 1. *This is a sample theorem. The run-in heading is set in bold, while the following text appears in italics. Definitions, lemmas, propositions, and corollaries are styled the same way.*

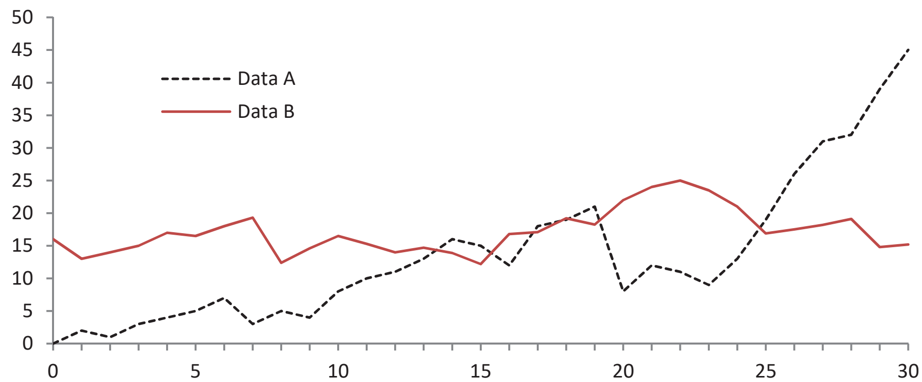

Fig. 1. A figure caption is always placed below the illustration. Please note that short captions are centered, while long ones are justified by the macro package automatically.

Proof. Proofs, examples, and remarks have the initial word in italics, while the following text appears in normal font.

For citations of references, we prefer the use of square brackets and consecutive numbers. Citations using labels or the author/year convention are also acceptable. The following bibliography provides a sample reference list with entries for journal articles [**?**], an LNCS chapter [**?**], a book [**?**], proceedings without editors [**?**], and a homepage [**?**]. Multiple citations are grouped [**?**,**?**,**?**], [**?**,**?**,**?**,**?**].

Acknowledgments. A bold run-in heading in small font size at the end of the paper is used for general acknowledgments, for example: This study was funded by X (grant number Y).

Disclosure of Interests. It is now necessary to declare any competing interests or to specifically state that the authors have no competing interests. Please place the statement with a bold run-in heading in small font size beneath the (optional) acknowledgments[4], for example: The authors have no competing interests to declare that are relevant to the content of this article. Or: Author A has received research grants from Company W. Author B has received a speaker honorarium from Company X and owns stock in Company Y. Author C is a member of committee Z.

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

[4] If EquinOCS, our proceedings submission system, is used, then the disclaimer can be provided directly in the system.

3. Friston, K.: The free-energy principle: a unified brain theory? Nature Reviews Neuroscience **11**(2), 127–138 (2010). https://doi.org/10.1038/nrn2787
4. 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