Literature Review

An Analysis of Active Inference and Reinforcement Learning Paradigms in Large, Partially Observable and Non-Stationary Environments.

Fraser Paterson

A review of the extant literature, pursuant to the requirements of the Degree: Bachelor of Science (Honours).

Supervisor: Dr Tim French
Department of Computer Science and Software Engineering
The University of Western Australia
24 April 2023

Contents

1 Introduction

1.1 Active Inference: An Overview

Active Inference is an emerging first-principles account of adaptive behaviour. Originating from Neuroscience: Friston, Kilner, and Harrison [13], Knill and Pouget [17] and Friston et al. [10], Active Inference is corollary of the "Free Energy Principle", which is a theoretical principle thought to plausibly offer a unified, constitutive account of brain function: Friston [12] and perhaps even of life itself: Friston [11].

Active Inference is increasingly making inroads into Machine Learning and Artificial Intelligence: Friston, Daunizeau, and Kiebel [15] and Millidge [22]. The theory is highly ambitious, as it purports to offer a fully unified account of action, perception and learning. The basic postulate of the theroy is that adaptive systems like living organisms will act to fulfill prior expectations or "preferences" which encode desierable states for the system. The system achieves this by means of a propabalistic generative model which

Active Inference is derived from the free energy principle...

Active inference is interesting because of its promise to be such a general method and indeed owing to a growing body of empirical research to suggest that free-energy minimization is what the brain is doing - I have citations for this. Since the brain is thought to be the seat of "natural" intelligence, evidence attesting to the brain's function as a "free-energy minimizing machine" must surely be of interest to we who are concerned with generating instances of intelligence, artificially.

This is fine, though try to avoid "grandiose" claims. A better approach is to lead with an explanation as to what it is that Active Inference sovles and perhaps how it is novel in a useful way.

Active Inference has typically only been implemented on relatively trivial problem instances, with a small number of states and/or actions, and in a discrete setting. T-maze, saccadic eye movements, examples... The goal of scaling up the method to problems with larger state and/or action spaces, such as in the continuous case, is an open problem, and one that would afford X, Y and Z valuable capabilities.

1.2 Active Inference: Key Concepts

Provide brief overview of key concepts and theories to be discussed in the review

Things that will some elaboration, or at least a definition:

1. Free Energy principle
2. Bayesian Inference
3. Variational Inference
4. Variational Free Energy
5. Expected Free Energy
6. Model-Based vs Model-Free Methods - not sure if this is necessary
7. Factor Graphs and Message Passing

8. Policy (Reinforcement Learning vs Active Inference framing)

9. Amortized Inference

There are too many here to do full justice in the manner you desire. Maintain a glossary of key terms, in addition.

1.3 Overview of Research Direction

Perhaps also reiterate the research question? - certainly later, after the body of the review.

The central aim of my proposed research topic is twofold. The first constituent aim is to investigate the relative merits/demerits of active inference as a real-world control and optimization strategy, against reinforcement learning baselines. I propose to investigate this by framing the question of real-world suitability in terms of the approach's ability to afford fast and reliable solutions in noisy, uncertain or partially observable environments. the subsequent active inference agents I develop will be compared to Reinforcement Learning baselines.

The second aim concerns the potential for "scaling up" Active Inference methods to continuous and/or higher-dimensional state-spaces. This is a natural corollary to the first aim, since if we are interested in the suitability of Active Inference as a real-world control and optimization technique, it is not enough to simply determine if it can favorably compare to an established method in the noisy case. Since the real-world tasks of interest are overwhelmingly characterized by a high degree of dimensionality, it is necessary to investigate the performance of Active inference in high-dimensional settings.

Thus are the central questions raised in this Thesis:

- Are Active Inference agents more robust to noisy observations and non-stationarity than a comparable RL baseline?

- What are the most promising avenues of investigation in the attempt to scale up active inference to larger problem instances?

2 Previous Work

This thesis aims to investigate two distinct approaches to the affordance of the use of active inference methods in larger problem instances than have been practical to tackle with the method. In addition, the investigation is also concerned with unearthing any potentially systematic advantages the method might have, compared to reinforcement learning, with respect to model robustness in the presence of observation noise.

Attempts to address these twin issues; scaling active inference to larger problem instances and investigating its robustness as compared with more standard methods, have already begun to appear in the literature, though these endeavors are still very much in their infancy.

Marković et al. [20] implemented an Active Inference agent for the multi-armed bandit problem, in the stationary and non-stationary case. In the stationary case, this agent Active did not perform as well as a special purpose, state-of-the-art

Bayesian UCB algorithm. However in the non-stationary case, the Active Inference agent outperformed the UCB agent. While this implementation was conducted over a small, discrete state-action space, the results plausibly suggest that Active Inference would be an effective means of robust inference and control in a higher-dimensional or continuous problem.

more on the robustness issues goes here...

2.1 Factor Graphs and Message Passing Methods

A particularly novel approach that has enjoyed some success as of late, casts the problem of inference as a species of message passing updates on a Forney factor graph: Cox, Laar, and Vries [6], Laar and Vries [19], Vries and Friston [33] and Bagaev and Vries [1].

In this framework, the agent's generative model is constructed in such a was as to instantiate a Forney or "Normal" factor graph: Forney [9]. Free Energy minimization is then cast as a process of message passing over this factor graph. Various message passing algorithms exist, such as Belief Propagation and Variational Message Passing. This message passing scheme greatly reduces the number of terms over which it is necessary to sum, when computing the approximate marginal and posterior distributions; affording much more efficient inference and a great potential for scaling up to larger state-action spaces. Indeed, this method does not make use of any approximation by means of a sampling procedure. Since this method relies upon a particular schedule of message-passing update rules on the underlying factor graph, all functions used need to be invertible and an inference is performed via a closed form update where the prior and likelihood distributions must be conjugate. The model passes around full distributions instead of mere samples. This results in a very fast and efficient implementation - when applicable, but the issue is that it is not a completely generic method as of yet, owing to the many assumptions as to the model structure just enumerated.

2.2 Sampling Based Approximation Methods

A more standard approach that has seen a comparatively greater deal of attention is that of using deep neural network function approximators to either parameterize the distributions of interest, or to afford an efficient means of sampling these distributions. This affords approximate inference. Indeed this "genre" of approach has already seen great success in scaling reinforcement learning methods to larger state-action spaces: Mnih et al. [24], Mnih et al. [25], other citations needed and available.

See: Ueltzhöffer [32], Millidge [22], Millidge [23], Tschantz et al. [30] and Çatal et al. [4].

Of particular interest are: Tschantz et al. [31], Çatal et al. [4] and Mazzaglia, Verbelen, and Dhoedt [21]. The former makes use of amortized inference, in the form of neural network function approximators to parameterize the relevant distributions. Free Energy minimization is then performed with respect to the function approximators. In addition, the free energy functional is amortized over the training data. This affords several advantages. For example, the number of parameters remains constant with respect to the size of the data and inference can be achieved via a single forward pass through the network. This contrasts with the iterative approach, where the VFE must be scored for every sample, individually. The resulting algorithm was able to explore a much greater proportion of the state space in a simple stationary environment, in comparison with two Reinforcement Learning baseline agents. In addition, the agent was able to learn to control the continuous inverted pendulum task with a far greater sample efficiency than the baseline agents. Although the approach offered in Tschantz et al. [31] is promising, its analysis was restricted in every case to fully observable environments. This potentially sold the implementaon short, since the partially observable domain is the more "natural" problem instance for which active inference was conceived as a solution strategy. Active Inference has a built-in drive to effect uncertainty reduction, this is not so with standard reinforcement learning, for which only ad-hoc strategies exist to afford the same sort of epistemic drive that exists in active inference. This is a salient point of departure between active inference and reinforcement learning, since both already implement strategies for realizing pragmatic value.

Pragmatic value is encoded as the sum of discounted reward across time, in the case of reinforcement learning. In the case of active inference, the drive to realize pragmatic value is afforded by the choice of action/s that realize the agent's prior preferences. The realization of prior preferences in active inference is analogous to maximizing the reward signal in reinforcement learning. However there is no analogous process for a drive to realize epistemic value, going from active inference to reinforcement learning, at least not without some aforementioned ad-hoc contrivance of the reward signal.

Lastly, the approach of Mazzaglia, Verbelen, and Dhoedt [21], implemented a contrastive method for their Active Inference agent, which significantly reduced the computational burden of learning the parameters for the generative model and planning future actions. This method performed substantially better than the usual, likelihood-based "reconstructive" means of implementing Active Inference and it was also computationally cheaper to train. Importantly, this method offered a unique way to afford increased model-robustness in the face of environmental distractors.

3 Gaps in Literature

Though there has been much focus on the implementation of active inference methods for small, discrete state-action spaces: Millidge [22], Da Costa et al. [7], Smith, Friston, and Whyte [27], Da Costa et al. [8] and Friston et al. [14]. The method is not currently viable for practical use in larger or continuous state-action spaces, for which it is necessary to plan future actions over some time horizon. Owing to the relatively small size of the state-action spaces in which active inference has historically been implemented, it has been possible to simply evaluate the expected free energy of all possible actions over the specified time horizon. This owes primarily to the issue of evaluating the expected free energy, which is the expectation of the Variational free energy evaluated for future actions over some time horizon: Koudahl, Buckley, and Vries [18] and Çatal et al. [4].

Unfortunately, enumerating all possible action-trajectories over the specified time horizon does not scale well to problems with larger state-action spaces and/or longer time horizons. Hence we can now specify exactly what it is that the problem of "scaling" is supposed to be.

Let S be a solution technique. Let P_1 and P_2 be problem instances of the same type. Let X and Y be the solution spaces for P_1 and P_2 (respectfully), where $|X| << |Y|$. Suppose the solution technique S affords an adequate solution to problem instance P_1, in the sense that the solution is both adequate for the task at hand and S found the solution in an adequate amount of time, consuming an acceptable amount of computational resources.

S will be is said to scale (or scale well) to problem instance P_2, if S can generate a solution to P_2, in an acceptable amount of time, while consuming an acceptable amount of computational resources. In other words, the cost associated with generating the solution to P_2 does not outweigh the utility of being able to generate the solution to P_2, S is a "viable" solution technique for instance P_2.

Evaluating all possible trajectories in a problem instance's state-action space, scales exponentially with the size of the state-action space: Millidge [22]. For large state-action spaces, evaluating all possible action trajectories quickly becomes an "unviable" solution technique.

4 Research Aims

I thought I'd reiterate the research questions here and aim to clarify what exactly it is that I seek to achieve.

5 Conclusion

Here I think I'll reiterate why this problem of scaling active inference is important at all and suggest potential implications for being able to make some headway in on this problem.

6 Overall Plan/Scaffold

This is simply the list of papers I intend on citing. I won't be able to cite all of them but I think it's best to write what I think I need to write and then cull.

Introductory/Context-Affording papers:

1. "The free-energy principle: a rough guide to the brain?": Friston [12]

2. "Reinforcement Learning: An Introduction": Sutton and Barto [28]

3. "Action and behavior: a free-energy formulation": Friston et al. [10]

4. "Mastering the game of Go without human knowledge": Silver et al. [26]

5. Issue scaling AIF "Active inference on discrete state-spaces: A synthesis": Da Costa et al. [7]

6. Issue scaling AIF "A step-by-step tutorial on active inference and its application to empirical data": Smith, Friston, and Whyte [27]

7. Issue scaling AIF "Applications of the FEP to ML and Neuroscience": Millidge [22]

8. Contemporary RL Method "Asynchronous Methods for Deep Reinforcement Learning": Mnih et al. [24]

9. Contemporary RL Method "Playing Atari with Deep Reinforcement Learning": Mnih et al. [25]

10. "Reinforcement Learning or Active Inference?": Friston, Daunizeau, and Kiebel [15]

11. "The FEP for action and perception a mathematical review": Buckley et al. [3]

12. "The Bayesian brain: the role of uncertainty in neural coding and computation": Knill and Pouget [17]

13. "An empirical evaluation of active inference in multi-armed bandits": Marković et al. [20]

Neural Network Approximators:

1. "Applications of the FEP to ML and Neuroscience": Millidge [22]
2. "Deep Active Inference": Ueltzhöffer [32]
3. "Deep Active Inference as Variational Policy Gradients": Millidge [23]
4. "Scaling Active Inference": Tschantz et al. [31]
5. "Reinforcement Learning Through Active Inference": Tschantz et al. [30]
6. "Bayesian Policy Selection Using Active Inference": Çatal et al. [4]
7. "Contrastive Active Inference": Mazzaglia, Verbelen, and Dhoedt [21]

Factor Graph and Message Passing Implementations

1. "Codes on graphs: normal realizations": Forney [9]
2. "Simulating Active Inference by Message Passing": Laar and Vries [19]
3. "Applications of the FEP to ML and Neuroscience": Millidge [22]
4. "A factor graph approach to automated design of Bayesian signal processing algorithms": Cox, Laar, and Vries [6]
5. "Reactive Message Passing for Scalable Bayesian Inference": Bagaev and Vries [1]
6. "A Factor Graph Description of Deep Temporal Active Inference": Vries and Friston [33]
7. "Deep Active Inference for Partially Observable MDPs": Himst and Lanillos [16]
8. "Bayesian policy selection using active inference": Çatal et al. [5]

Finally, address the specific papers on Scaling that I'll use:

1. Sampling/Neural Networks "Scaling Active Inference": Tschantz et al. [31]
2. Sampling?/neural nets? " "Contrastive Active Inference": Mazzaglia, Verbelen, and Dhoedt [21]

Other Papers:

1. Paper: "The FEP for action and perception a mathematical review": Buckley et al. [3]
2. paper: "Simulating Active Inference by Message Passing": Laar and Vries [19]
3. Paper: "a practical tutorial on Variational Bayes": Tran, Nguyen, and Dao [29]
4. Paper: "action and behavior, a free energy formulation": Friston et al. [10]
5. paper: "A tutorial on the free-energy framework for modelling perception and learning": Bogacz [2]
6. Paper: "A step-by-step tutorial on active inference and its application to empirical data": Smith, Friston, and Whyte [27]
7. Paper: "Scaling Active Inference": Tschantz et al. [31]
8. paper: The Cape Town AIF/RL Honours Thesis - not sure if published/if appropriate
9. PhD Thesis: "Applications of the FEP to ML and Neuroscience": Millidge [22]

References

[1] Dmitry Bagaev and Bert de Vries. "Reactive Message Passing for Scalable Bayesian Inference". In: *CoRR* abs/2112.13251 (2021). arXiv: 2112.13251. URL: https://arxiv.org/abs/2112.13251.

[2] Rafal Bogacz. "A tutorial on the free-energy framework for modelling perception and learning". In: *Journal of Mathematical Psychology* 76 (2017). Model-based Cognitive Neuroscience, pp. 198–211. ISSN: 0022-2496. DOI: https://doi.org/10.1016/j.jmp.2015.11.003. URL: https://www.sciencedirect.com/science/article/pii/S0022249615000759.

[3] Christopher L. Buckley et al. "The free energy principle for action and perception: A mathematical review". In: *Journal of Mathematical Psychology* 81 (2017), pp. 55–79. ISSN: 0022-2496. DOI: https://doi.org/10.1016/j.jmp.2017.09.004. URL: https://www.sciencedirect.com/science/article/pii/S0022249617300962.

[4] Ozan Çatal et al. *Bayesian policy selection using active inference*. 2019. arXiv: 1904.08149 [cs.LG].

[5] Ozan Çatal et al. *Bayesian policy selection using active inference*. 2019. arXiv: 1904.08149 [cs.LG].

[6] Marco Cox, Thijs van de Laar, and Bert de Vries. "A factor graph approach to automated design of Bayesian signal processing algorithms". In: *International Journal of Approximate Reasoning* 104 (Jan. 2019), pp. 185–204. DOI: 10.1016/j.ijar.2018.11.002.

[7] Lancelot Da Costa et al. "Active inference on discrete state-spaces: A synthesis". In: *Journal of Mathematical Psychology* 99 (2020), p. 102447. ISSN: 0022-2496. DOI: https://doi.org/10.1016/j.jmp.2020.102447. URL: https://www.sciencedirect.com/science/article/pii/S0022249620300857.

[8] Lancelot Da Costa et al. *The relationship between dynamic programming and active inference: the discrete, finite-horizon case*. Sept. 2020.

[9] G.D. Forney. "Codes on graphs: normal realizations". In: *IEEE Transactions on Information Theory* 47.2 (2001), pp. 520–548. DOI: 10.1109/18.910573.

[10] K Friston et al. "Action and behavior: a free-energy formulation". In: *Biological cybernetics* 102.3 (2010), pp. 227–260. DOI: 10.1007/s00422-010-0364-z. URL: https://doi.org/10.1007/s00422-010-0364-z.

[11] Karl Friston. "Life as we know it". In: *Journal of The Royal Society Interface* 10.86 (2013), p. 20130475. DOI: 10.1098/rsif.2013.0475. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rsif.2013.0475. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2013.0475.

[12] Karl Friston. *The free-energy principle: a rough guide to the brain?* 2009. URL: https://doi.org/10.1016/j.tics.2009.04.005.

[13] Karl Friston, James Kilner, and Lee Harrison. "A free energy principle for the brain". In: *Journal of Physiology-Paris* 100.1 (2006). Theoretical and Computational Neuroscience: Understanding Brain Functions, pp. 70–87. ISSN: 0928-4257. DOI: https://doi.org/10.1016/j.jphysparis.2006.10.001. URL: https://www.sciencedirect.com/science/article/pii/S092842570600060X.

[14] Karl Friston et al. "Active inference and epistemic value". In: *Cognitive Neuroscience* 6.4 (2015). PMID: 25689102, pp. 187–214. DOI: 10.1080/17588928.2015.1020053. eprint: https://doi.org/10.1080/17588928.2015.1020053. URL: https://doi.org/10.1080/17588928.2015.1020053.

[15] Karl J. Friston, Jean Daunizeau, and Stefan J. Kiebel. "Reinforcement Learning or Active Inference?" In: *PLOS ONE* 4.7 (July 2009), pp. 1–13. DOI: 10.1371/journal.pone.0006421. URL: https://doi.org/10.1371/journal.pone.0006421.

[16] Otto van der Himst and Pablo Lanillos. "Deep Active Inference for Partially Observable MDPs". In: *Active Inference*. Ed. by Tim Verbelen et al. Cham: Springer International Publishing, 2020, pp. 61–71. ISBN: 978-3-030-64919-7.

[17] David C. Knill and Alexandre Pouget. "The Bayesian brain: the role of uncertainty in neural coding and computation". In: *Trends in Neurosciences* 27.12 (2004), pp. 712–719. ISSN: 0166-2236. DOI: https://doi.org/10.1016/j.tins.2004.10.007. URL: https://www.sciencedirect.com/science/article/pii/S0166223604003352.

[18] Magnus Koudahl, Christopher L. Buckley, and Bert de Vries. "A Message Passing Perspective on Planning Under Active Inference". In: *Active Inference*. Ed. by Christopher L. Buckley et al. Cham: Springer Nature Switzerland, 2023, pp. 319–327. ISBN: 978-3-031-28719-0.

[19] Thijs W. van de Laar and Bert de Vries. "Simulating Active Inference Processes by Message Passing". In: *Frontiers in Robotics and AI* 6 (2019). ISSN: 2296-9144. DOI: 10.3389/frobt.2019.00020. URL: https://www.frontiersin.org/articles/10.3389/frobt.2019.00020.

[20] Dimitrije Marković et al. "An empirical evaluation of active inference in multi-armed bandits". In: *Neural Networks* 144 (2021), pp. 229–246. ISSN: 0893-6080. DOI: https://doi.org/10.1016/j.neunet.2021.08.018. URL: https://www.sciencedirect.com/science/article/pii/S0893608021003233.

[21] Pietro Mazzaglia, Tim Verbelen, and Bart Dhoedt. "Contrastive Active Inference". In: *Advances in Neural Information Processing Systems*. Ed. by M. Ranzato et al. Vol. 34. Curran Associates, Inc., 2021, pp. 13870–13882. URL: https://proceedings.neurips.cc/paper_files/paper/2021/file/73c730319cf839f143bf40954448ce39-Paper.pdf.

[22] Beren Millidge. "Applications of the free energy principle to machine learning and neuroscience". PhD thesis. University of Edinburgh, 2021. URL: https://era.ed.ac.uk/handle/1842/38235.

[23] Beren Millidge. "Deep Active Inference as Variational Policy Gradients". In: (July 2019). URL: https://arxiv.org/pdf/1907.03876.pdf.

[24] Volodymyr Mnih et al. *Asynchronous Methods for Deep Reinforcement Learning*. 2016. arXiv: 1602.01783 [cs.LG].

[25] Volodymyr Mnih et al. *Playing Atari with Deep Reinforcement Learning*. 2013. arXiv: 1312.5602 [cs.LG].

[26] David Silver et al. "Mastering the game of Go without human knowledge". In: *Nature* 550.7676 (Oct. 2017), pp. 354–359. ISSN: 1476-4687. DOI: 10.1038/nature24270. URL: https://doi.org/10.1038/nature24270.

[27] Ryan Smith, Karl J. Friston, and Christopher J. Whyte. "A step-by-step tutorial on active inference and its application to empirical data". In: *Journal of Mathematical Psychology* 107 (2022), p. 102632. ISSN: 0022-2496. DOI: https://doi.org/10.1016/j.jmp.2021.102632. URL: https://www.sciencedirect.com/science/article/pii/S0022249621000973.

[28] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. 2nd ed. Cambridge, Massachusetts: MIT Press, 2018.

[29] Minh-Ngoc Tran, Trong-Nghia Nguyen, and Viet-Hung Dao. *A practical tutorial on Variational Bayes*. 2021. arXiv: 2103.01327 [stat.CO].

[30] Alexander Tschantz et al. *Reinforcement Learning through Active Inference*. 2020. arXiv: 2002.12636 [cs.LG].

[31] Alexander Tschantz et al. "Scaling active inference". In: *CoRR* abs/1911.10601 (2019). arXiv: 1911.10601. URL: http://arxiv.org/abs/1911.10601.

[32] Kai Ueltzhöffer. "Deep active inference". In: *Biological Cybernetics* 112.6 (Oct. 2018), pp. 547–573. DOI: 10.1007/s00422-018-0785-7.

[33] Bert de Vries and Karl J. Friston. "A Factor Graph Description of Deep Temporal Active Inference". In: *Frontiers in Computational Neuroscience* 11 (2017). ISSN: 1662-5188. DOI: 10.3389/fncom.2017.00095. URL: https://www.frontiersin.org/articles/10.3389/fncom.2017.00095.