

Toward Scalable and Robust Agent Based Methods.

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

Fraser Paterson

A Thesis proposal, pursuant to the requirements of the
Degree: Bachelor of Science (Honours).



THE UNIVERSITY OF
WESTERN
AUSTRALIA

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

Contents

0.1	Introduction	3
0.1.1	Motivation and Background	3
0.1.2	Active Inference: An Overview	4
0.2	Previous Work	5
0.3	Research Objectives	5
0.3.1	Central Research Question	5
0.4	Methods	5
0.4.1	Techniques to be Investigated	5
0.4.2	Metrics	5
0.4.3	Software and Data	5
0.4.4	Simplifying Assumptions	6
0.5	Schedule	6

0.1 Introduction

0.1.1 Motivation and Background

Many real-world problems are characterised by high degrees of noise, ill-definedness and uncertainty. This uncertainty assumes myriad forms, whether in the clarity of the observations one can solicit from the system of interest (signal detection), or in the confidence of an inference as to the system-parameter values, that best account for the solicited observation (state-estimation). These tasks are only further complicated by a very common constraint on any candidate solution technique: partial-observability, which is an obvious source of uncertainty. Indeed, partial-observability is overwhelmingly characteristic of many “difficult” real-world systems, [examples and citations?](#).

It is the prerogative of the Engineer to make simplifying assumptions about the kinds and sources of uncertainty which beset the task of modeling, controlling and/or predicting the behaviour of a system. A “good” simplification (model) of a complicated system, affords some relevant trade-off between the computational resources necessary to predict and control the entities of interest, as against the material and temporal constraints necessitated by the nature of the problem. This trade-off is inherent to cognitive agency itself; it is to be in the “Finitary Predicament”, see: Cherniak [1]. For example, a “pedestrian-detection” algorithm on a self-driving car, plausibly needs to put a premium on speed since temporal constraints are quite salient in this case. An algorithm for optimising the trajectory of a spacecraft mission 5 years from now faces the same trade-off, but under near-inverse constraints. In this case, it is plausible to suppose that there is not such a high premium on time. In any case, a perennial problem remains, that of obtaining accurate *enough* observations and performing an optimal *enough* action - at the optimal *enough* location and time - for resolving the maximal amount of uncertainty about the dynamics one wishes to predict and/or control, while also minimising the costs associated with planning and performing these actions.

The faculty by which complex biological organisms achieve these ends is by means of their intelligence. A relatively non-controversial definition of *Intelligence* is “general problem-solving ability”. The ongoing project of defining, detecting and creating instances of intelligence is one that takes a startling variety of forms. Psychologists speak of “intelligent behaviour”, whereas to the Neuroscientist, intelligence is studied in terms of “patterns of neuronal activation”. Yet again, the computer scientist traditionally accounts for intelligence as a kind of “information processing”. Insofar as these disciplines might share a common project, it may plausibly be supposed as the attempt to understand the nature and function of intelligence. Several philosophical approaches to this project abound - as attested by the variety of disciplines it spans. Again, the Psychologist studies *behaviour*, the neuroscientist studies *the brain* but the computer scientist has an altogether distinct modality of investigation. This “investigative modality” consists in the attempt to build, create and/or instantiate an *instance* of intelligence, hence “Artificial” Intelligence. The hope is that if we are successful enough at reproducing the effects of natural intelligence, that this will count as a good reason to suppose that the nature and function of our artificial intelligence, is constitutive of the nature and function

of “Natural Intelligence”. Thus the project of Artificial Intelligence consists in the attempt to analyse, formalise and mechanise Natural Intelligence.

An early and perhaps paradigm case of an attempt to so analyse, formalise and mechanise intelligence was Newell and Simon’s General Problem Solver (GPS): Newell and Simon [3]. Newell and Simon attempted to specify *what it was to have a problem at all*. For them, a *problem specification* consisted in a specification of the goal state, the starting state, a set of operators that can transform one state into another and lastly, a set of path-constraints that any trajectory must satisfy in order to be considered a valid trajectory. The set of all trajectories from the start to the goal state is the “search space” and hence a problem solution consists in finding the optimal trajectory through this search space.

- Intelligence as “general problem Solving ability”. Newell and Simon and the GPS.
- The justification for “biologically inspired” A.I - we can do it very well so let’s pattern AI on us...
- Insight into the structure and function of AI might come from the attempt to build it: analyse, formalise and mechanise
- Newell and Simon, the GPS - as an example of analysis, formalisation and mechninisation.
- Combinatorial explosion, the necessity for biased attention.
- Model based AI as a means for this.
- dealing with combinatorial explosion in model based methods.
- optimal behaviour as the satisfcation of imperatives: “to maximise reward”, “to minimise uncertainty”, “to seek information”.
- The conflict between imperatives: explore-exploit trade-off?
- AIF and RL as aproaches to afford intelligent agency.
- Scalibility issues. Robusteness Issues (the money-shot).

0.1.2 Active Inference: An Overview

- The FEP
- Variational inference
- Lightening quick exposition of the general theroy

AIF is an agent-based method which makes use of a probabilistic, generative model over sensory observations, environmental states and various parameters necessary to encode the Variational Free Energy (VFE). The ontological basis of Active Inference is broadly similar to various, older agent based methods such as Reinfrcement Learning: Cite RL. An AIF agent performs approximate bayesian inference

over a hidden state-space to optimise/control some external process, in the service of some goal, where this goal is encoded as a Bayesian prior belief (a distribution over observations).

Due to the fact that exact inference is usually intractable, AIF introduces an approximate posterior distribution over sensory states: $q(\phi)$, where we denote the true posterior as $p(\phi|e)$, where ϕ is sensory data of one kind or another and e is an environmental variable in the hidden-state space. We can then construct a “measure” - in the colloquial sense - of distance between these two densities, as the kullback-leibler divergence between each density. Assuming the discrete case:

A Simple Example

Manuel’s minimal agent

0.2 Previous Work

- RL and AIF
- AIF on a Factor Graph and Message-Passing
- Manuel’s paper
- Contrastive Methods paper
- Latent Space paper
- Marković et al. [2]

0.3 Research Objectives

0.3.1 Central Research Question

The principle aim of this research thesis will be to investigate the relative advantages and disadvantages of AIF methods, as against more “traditional” methods of inference and control, such as the well-known Linear Quadratic Gaussian (LGQ), and Reinforcement Learning (RL) algorithms. Of specific interest, will be the question as to the performance of AIF as against the above methods, in the non-stationary control and inference environments.

The natural inclusion of a generative model and the built-in epistemic imperatives - toward the aim of uncertainty reduction - in AIF, make it highly plausible that this method will be better able to deal with non-stationary environments, dynamic constraint changes, noise and other such sources of uncertainty. Indeed, just as much has been shown in Marković et al. [2], where AIF performed better than a strong Bayesian UCB algorithm in a non-stationary multi-armed bandit problem.

0.4 Methods

0.4.1 Techniques to be Investigated

0.4.2 Metrics

0.4.3 Software and Data

- RxInfer.jl, see the mountaincar implementation: <https://biaslab.github.io/RxInfer.jl>
- OpenAI gym and Atari Environments
- DeepMind Control Suite: <https://www.deepmind.com/open-source/deepmind-control-suite>

0.4.4 Simplifying Assumptions

Factor Graphs

Laplace Approximation

Mean-Field Assumption

0.5 Schedule

Bibliography

- [1] Christopher Cherniak. *Minimal Rationality*. MIT Press, 1986.
- [2] 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>.
- [3] Allan Newell and Herbert Simon. *Human problem solving*. Prentice-Hall, 1972.