

University of Innsbruck
Faculty of Mathematics, Computer Science and Physics
Institute for Theoretical Physics



Bachelor Thesis
submitted for the degree of
Bachelor of Science

**An information-theoretical approach to
internal models in a Partially Observable
Markov Decision Process**

by
Lukas Prader
Matriculation Nr.: 12115058
SE Seminar with Bachelor Thesis

Submission Date: 15th June 2024
Supervisors: Alexander Vining, Hans J Briegel

Abstract

Lorem ipsum

Contents

1	Introduction	1
2	Methods	3
3	Results	4
4	Conclusion	6
5	Acknowledgements	6

1 Introduction

A large reason for success in many, if not all scientific disciplines has been the adoption of a reductionist perspective on phenomena. Widely adopted and successful, this hypothesis assumes that all processes in our universe are in the end governed by a set of fundamental laws, which can be used to describe any higher order of phenomena as well. Especially in physics, the idea of a final, unified theory of everything has been seen as the ultimate goal of the discipline for centuries.

Yet, in many disciplines it has been shown that there are so-called emergent phenomena, which are not easily explained by lower-level fundamental laws, requiring additional concepts to accurately explain them (Anderson 1972). Exactly how these complex phenomena can emerge from the set of currently known fundamental laws is an ongoing field of research, with interdisciplinary approaches taking from the fields of physics, chemistry, biology, psychology, philosophy, computer science and others.

Especially the rise of artificial intelligence in recent years has produced new research trying to create complex models able to perform intelligent tasks. Research into so-called complex adaptive systems is also connected to research in biology, trying to understand the emergence of intelligent and adaptive behaviour in biological organisms. Gaining insight into the mechanisms which enable biological systems to exhibit complex behaviour can in turn be used to improve the ways in which we attempt to create systems capable of these behaviours ourselves.

Complex system research can thus provide valuable insights into the emergence of intelligent behaviour, which also includes processes connected to adaptive behaviour and learning in animals. Works by the likes of Pavlov and Skinner have probed into the mechanisms that influence behavioural patterns in animals (Pavlov 1906; Skinner 1957). Exactly how conditionable behaviour like this can emerge just from interaction with the environment is still not fully understood.

Information theory has been fundamental in furthering our understanding of such complex processes. With the work of Shannon (Shannon 1948) as a basis, modern information theory has enabled researchers to quantify correlations and changes in complexity of dynamic systems. It has successfully been applied to many examples in behavioural biology, such as looking into the collective behaviour of ants, information exchange in slime molds and group behaviour in bat populations (Kim et al. 2021).

Information theory can also be used to examine how biological agents, such as animals, acquire an understanding of their surroundings in order to then act in response, a so-called internal model of their environment. Current approaches to create agents able to learn certain behaviours rely on trial and error to find the amount of param-

eters necessary to explain the complexity of a given environment, especially if only limited information about an environment is available to the agent. Real life organisms do not seem to be limited as much, still being able to infer information about their environment even with limited information to their disposal.

Crutchfield and Feldman (Crutchfield et al. 2003) have proposed information theoretical quantities able to explain processes of inference for complex systems. They can quantify the amount of "synchronization" a system has achieved in comparison to a different system (like an environment), which it can interact with. This framework can provide tools which may enable an agent to autonomously modify the current internal model in order to more accurately reflect the processes of the observed environment.

This thesis aims to apply some of the quantities proposed by Crutchfield and Feldman to a simple example of an agent acting with partial information of its environment, showing the capabilities of this framework to enable evaluation of the current internal model and subsequent improvements to accurately reflect the environment, even with only partial information available to the agent.

2 Methods

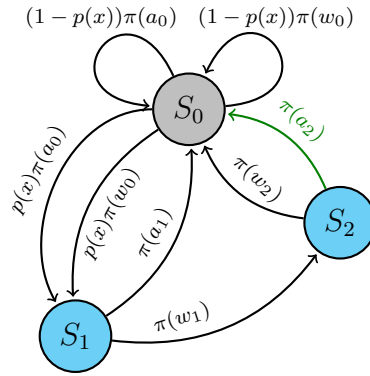


Figure 1: Graph showing the delayed action MDP

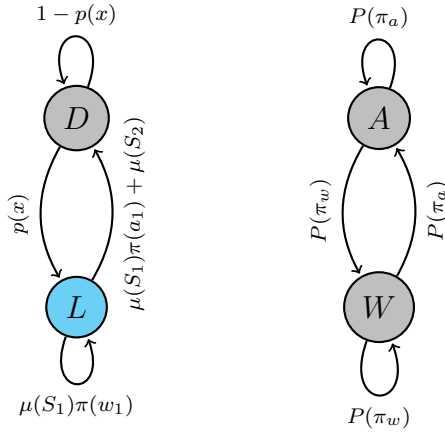


Figure 2: Reduced models for Observation of light (left) and action (right).

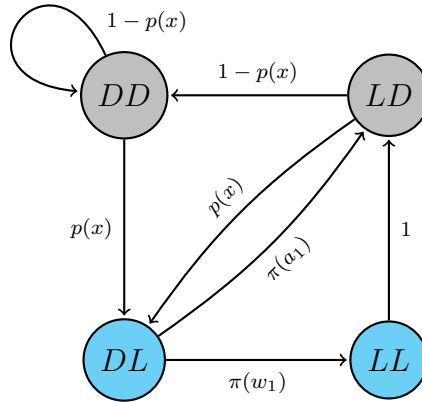


Figure 3: Internal model with sufficient complexity.

3 Results

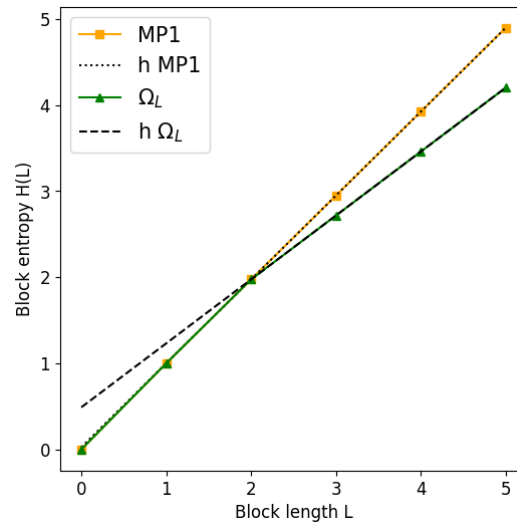


Figure 4: Plot showing the difference between MP1 and the MDP.

Talk about behaviour correlating to order 2 of process, in agreement with Crutchfield. Also do the same plot for MP2 and action as well, showing hidden Markov convergence.

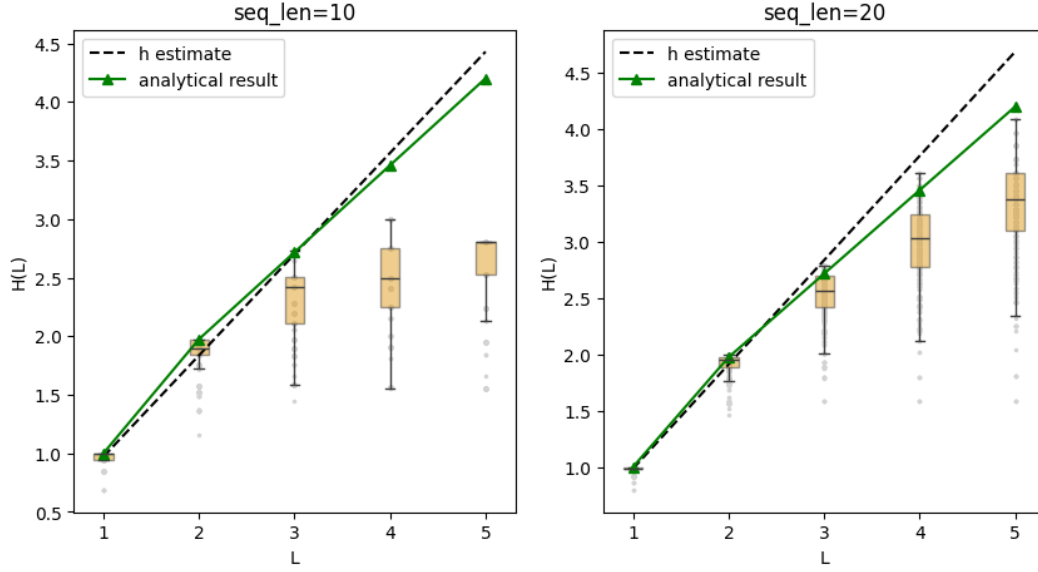


Figure 5: Plot showing the bias and variance of the plug in estimator for block Entropy.

$$\hat{H}(s^L) = - \sum_i \frac{n_i}{N} \log_2 \frac{n_i}{N}$$

(Lesne et al.) (10.1103/PhysRevE.79.046208) Entropy estimation of short (time correlated) sequences Upper bound on block length given a sequence of observations, in order to have good estimates:

$$n \leq \frac{N h_\mu}{\ln(k)}$$

with word length n , length of observation sequence N and number of symbols k . They propose first estimating the entropy rate using Lempel-Ziv complexity(iterative algorithm moving through sequence), then setting the sequence length/word length accordingly.

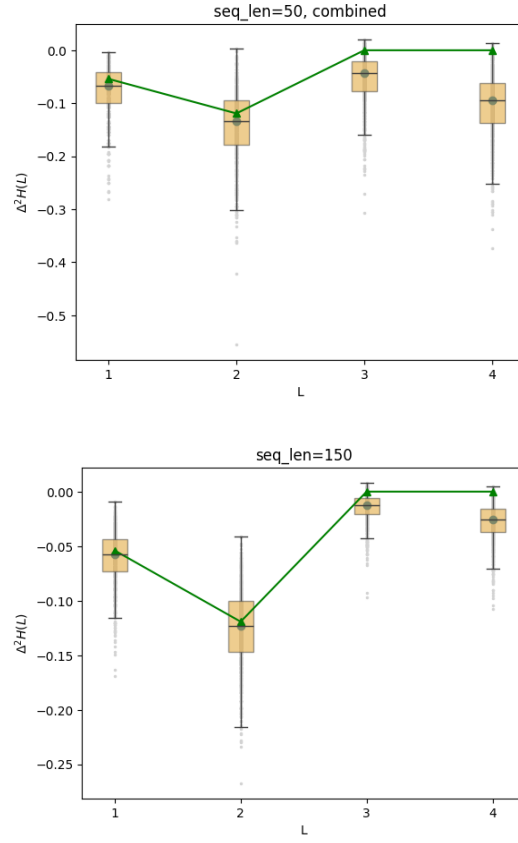
$$\hat{L}_0 = \frac{\mathcal{N}_w \ln(N)}{N}, \quad \hat{L} = \frac{\mathcal{N}_w [1 + \log_k \mathcal{N}_w]}{N}, \quad \lim_{n \rightarrow \infty} \hat{L} = \frac{h_\mu}{\ln(k)}$$

With N the length of the observation sequence and \mathcal{N}_w parsed words from the Lempel-Ziv algorithm. For the plug in estimator, error bars are computable, meaning computable confidence intervals.

Larson et al. (10.1016/j.procs.2011.04.172) use

$$\mathcal{L} h_\mu \geq L |A|^L \ln |A|$$

solving for L given \mathcal{L} returns $W(\mathcal{L} h_\mu) / \ln |A|$ (mathematica), with the Lambert W function. (gives better estimate)



deviation from analytical: [-0.00365034217876947 -0.00365907607011157 -0.0124851947847702 -0.0251811425394552]

4 Conclusion

5 Acknowledgements

Declaration of Authorship

I hereby solemnly declare, by my own signature, that I have independently authored the presented work and have not used any sources or aids other than those indicated. All passages taken verbatim or in content from the specified sources are identified as such.

I consent to the archiving of this Bachelor thesis.

Innsbruck, 15th June 2024

Lukas Prader



References

- Anderson, Philip W (1972). 'More Is Different: Broken symmetry and the nature of the hierarchical structure of science.' In: *Science* 177.4047, pp. 393–396. URL: <https://www.jstor.org/stable/pdf/1734697.pdf>.
- Crutchfield, James P and David P Feldman (2003). 'Regularities unseen, randomness observed: Levels of entropy convergence'. In: *Chaos: An Interdisciplinary Journal of Non-linear Science* 13.1, pp. 25–54.
- Kim, Hyunju et al. (2021). 'Informational architecture across non-living and living collectives'. In: *Theory in Biosciences*, pp. 1–17. DOI: 10.1007/s12064-020-00331-5.
- Pavlov, Ivan P (1906). 'The scientific investigation of the psychical faculties or processes in the higher animals'. In: *Science* 24.620, pp. 613–619. DOI: 10.1126/science.24.620.613.
- Shannon, Claude Elwood (1948). 'A mathematical theory of communication'. In: *The Bell system technical journal* 27.3, pp. 379–423. URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6773024>.
- Skinner, Burrhus Frederic (1957). 'The experimental analysis of behavior'. In: *American scientist* 45.4, pp. 343–371. URL: <https://www.jstor.org/stable/pdf/27826953.pdf>.