UNIVERSITY OF TARTU
Institute of Computer Science
Computer Science Curriculum

Sten Sootla

Analysing information distribution in complex systems

Bachelor's Thesis (9 ECTS)

Supervisor: Raul Vicente Zafra, PhD
Supervisor: Dirk Oliver Theis, PhD

Tartu 2017

Contents

1 Introduction **3**

2 Background **4**
2.1 Classical information theory 4
2.1.1 Entropy 4
2.1.2 Joint and Entropy 4
2.1.3 Kullback-Leibler distance 4
2.1.4 Mutual information 4
2.2 Partial information decomposition 4
2.3 Numerical estimator 4

3 Elementary cellular automata **5**
3.1 Problem description 5
3.2 Related work 5
3.3 Experimental setup 5
3.4 Results 5
3.5 Discussion 5

4 Ising model **6**
4.1 Problem description 6
4.2 Related work 6
4.3 Experimental setup 6
4.4 Results 6
4.5 Discussion 6

5 Neural networks **7**
5.1 Problem description 7
5.2 Related work 7
5.3 Experimental setup 7
5.4 Results 7
5.5 Discussion 7

6 Conclusion **8**

1 Introduction

Complex systems

In Chapter 1, the basics of classical information theory and partial information decomposition are covered. The chapter ends with an overview of the numerical estimator for PID.

The subsequent 3 chapters each introduce a specific complex system and the results of measuring information distribution in it.

Chapter 2 - Elementary Cellular Automata Chapter 3 - Ising model Chapter 4 - Artificial Neural Networks

In the final, concluding chapter, a summary of the contributions of this thesis is given, alongside suggestions for further work.

2 Background

2.1 Classical information theory

2.1.1 Entropy

Let X be a discrete random variable with a set of values from the set $\{x_1, x_2, ..., x_n\}$ and probability mass function $p(x_i) = P\{X = x_i\}$ $(i = 1, ..., n)$. Shannon [Sha48] defines the *entropy* of X as follows:

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i).$$

2.1.2 Joint and Entropy

2.1.3 Kullback-Leibler distance

2.1.4 Mutual information

2.2 Partial information decomposition

2.3 Numerical estimator

3 Elementary cellular automata

3.1 Problem description

3.2 Related work

3.3 Experimental setup

3.4 Results

3.5 Discussion

4 Ising model

4.1 Problem description

4.2 Related work

4.3 Experimental setup

4.4 Results

4.5 Discussion

5 Neural networks

5.1 Problem description

5.2 Related work

5.3 Experimental setup

5.4 Results

5.5 Discussion

6 Conclusion

References

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

Non-exclusive licence to reproduce thesis and make thesis public

I, Sten Sootla (date of birth: 17th of January 1995),

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

1.1 reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

1.2 make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

Analysing information distribution in complex systems

supervised by Raul Vicente Zafra and Dirk Oliver Theis

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, dd.mm.yyyy