Abstract

Write your abstract here...

Sammendrag

Skriv sammendrag her...

Preface

Write your preface here...

Acknowledgements (optional)

Table of Contents

List of Tables

List of Figures

Abbreviations

Symbol = definition

Chapter 1

Introduction

1.1 Background, Motivation and Problem Outline

1.2 Research Context

The research was conducted as my Master's thesis at the department of Computer and Information Science at the Norwegian University of Science and Technology. The research task was formulated by Odd Erik Gundersen, my supervisor, and is a continuation of previous work by Gundersen (2015) presented at 3DOR2015[1] Gundersen (2015).

1.3 Hypothesis, Objectives and Research Questions

Underlying this thesis is the hypothesis that; *the documentation provided in experimental publications at AI conferences is not good enough to consider the experiments reproducible.*

Objective 1 *Evaluate the reproducibility of accepted papers to AI conferences.*

RQ1 What is the state of reproducibility at AI conferences?

Objective 2 *Recommend practices that could be adopted to aid the reproducibility of conference papers.*

RQ2 What is generally missing from AI papers to support reproducibility?

RQ3 What can ease the documentation of missing information from conference papers?

[1] `http://vc.ee.duth.gr/3DOR2015/`

1.4 Research Approach

1.5 Research Contributions

C1: A survey of experimental research papers from AI conferences.

C2: An indication of the state of reproducibility at AI conferences.

C3: An approach to measure the reproducibility of AI conference papers.

1.6 Thesis Structure

1.7 Equations

To write an equation

```
\begin{eqnarray}\label{eq1}
F = m \times a
\end{eqnarray}
```

This will produce

$$F = m \times a \tag{1.1}$$

To refer to the equation

```
\eqref{eq1}
```

This will produce (1.1).

1.8 Figures

To create a figure

```
\begin{figure}[h!]
  \centering
    \includegraphics[width=0.5\textwidth]{fig/pikachu}
  \caption{Pikachu.}
\label{fig1}
\end{figure}
```

To refer to the figure

```
\textbf{Fig. \ref{fig1}}
```

This will produce **Fig. 1.1**

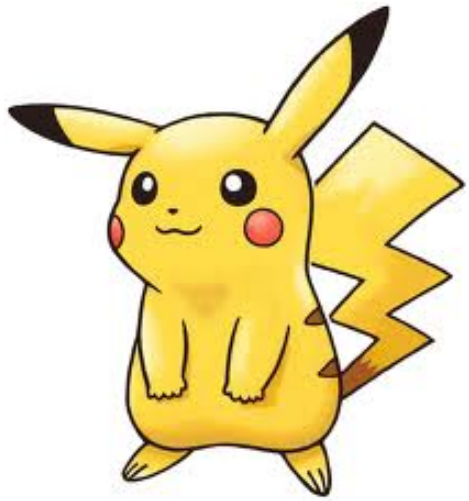

Figure 1.1: Pikachu.

1.9 References

To cite references

```
\cite{1,2,3}
```

or

```
\citep{1,2,3}
```

This will produce: Sarma and Chen (2008); Brouwer and Jansen (2004); Muskat (1937) or (Sarma and Chen, 2008; Brouwer and Jansen, 2004; Muskat, 1937), respectively.

1.10 Tables

To create a table

```
\begin{table}[!h]
\begin{center}
    \begin{tabular}{ | l | l | l | l |}
    \hline
    \textbf{No.} & \textbf{Data 1} & \textbf{Data 2} \\ \hline
     1 & a1 & b1 \\ \hline
     2 & a2 & b2 \\ \hline
    \end{tabular}
\end{center}
\caption{Table 1.}
```

```
\label{Tab1}
\end{table}
```

This will produce

No.	Data 1	Data 2
1	a1	b1
2	a2	b2

Table 1.1: Table 1.

To refer to the table

```
\textbf{Table. \ref{Tab1}}
```

This will produce **Table. 1.1**.

Chapter 2

Literature Review

2.1 Reproducibility Terminology

2.2 Best Practices for Reproducibility

2.3 Reproducibility in Practice

2.4 Observations

Chapter 3

Research Method

3.1 Litterature Survey Design

Advantages of doing a survey - Can be replicated on similar documents or on original documents provided the method is shared and documents are accessible - Can produce a lot of data at a low cost, in a relatively short time compared to attempting full replications of experiments - Allows a larger sample population due to the shorter time necessary to evaluate a paper

Disadvantages - The depth is restricted, does not provide detail on the research topic - Focuses on what can be counted and measured, other aspects may be overlooked

3.1.1 Data requirements

Want to investigate the reproducibility of experiments published in papers at AI conferences.

/* Explain variables and relate to best practices!!! */

Directly topic related : Is source code or data open for the experiment and method? Is the method documented? Is the experiment documented? etc.

Indirectly topic related Research transparency (hypothesis, predictions...) Author affiliation (uni/industry/both) Novel research? Conference view on supplementary material Theoretical / Experimental research

Possible analysis patterns

1. reproducibility related to author affiliation
2. reproducibility related to conference view on supplementary material?
3. reproducibility related to publishing year (improvement over time?)

3.1.2 Data generation method

Documents, conference papers. (Ch. 16) - Existing conference papers.

Advantages: - easy to obtain, accessible and are obtained unobtrusively - allows later longitudinal studies - other researchers can check and scrutinize the research based on original material

Disadvantages: -

Sampling frame: accepted papers at IJCAI-13, -16 and AAAI-14 and -16 (can be seen in repo files for sample generation) Sampling technique: probabilistic random sampling of each conference separately. ”Probability sampling, as its name suggests, means that the sample has been chosen because the researcher believes that there is a high probability that the sample of respondents (or events) chosen are representative of the overall population being studied. That is, they form a representative cross-section of the overall population.” Oates p.96 Discuss representativeness of sample method.

Sample size: 100 for each conference, restricts the necessary time to conduct the survey while still providing informative accuracy ranges when considering previous research (cite?)

Conference	Population Size	Sample Size	Confidence Interval
AAAI 2014	398	100	8.49
AAAI 2016	548	100	8.87
IJCAI 2013	413	100	8.54
IJCAI 2016	551	100	8.87
Combined	1910	400	4.36

Table 3.1: Confidence intervals of survey sample populations given a 50/50 yes/no split with confidence level of 95%. (`https://www.surveysystem.com/sscalc.htm`)

3.2 Evaluation Procedure

- Step by step ’instructions’ - Sampling documentation - Evaluation documentation - Example evaluations (variable X: ”Exhibit A” covers, ”Exhibit B” is not enough)

3.3 Limitations of the Survey

- Evaluation bias (modification of variables) - Sample inconsistency for IJCAI-13 (50 papers) - Not an actual attempt at reproducing experiments, researcher’s view that discussion of a variable is missing?

Chapter 4

Results and Analysis

Chapter 5

Discussion

- Anonymous publication of source code and data along with papers (blind review)

Chapter 6

Conclusion

Bibliography

Brouwer, D. R., Jansen, J. D., 2004. Dynamic optimization of waterflooding with smart wells using optimal control theory. SPE Journal 9 (4), 391–402.

Gundersen, O. E., 2015. Towards Scientific Benchmarks: On Increasing the Credibility of Benchmarks. In: Pratikakis, I., Spagnuolo, M., Theoharis, T., Gool, L. V., Veltkamp, R. (Eds.), Eurographics Workshop on 3D Object Retrieval. The Eurographics Association.

Muskat, M., 1937. Flow of Homogeneous Fluids. McGraw Hill.

Sarma, P., Chen, W. H., 2008. Applications of optimal control theory for efficient production optimization of realistic reservoirs. In: Proceedings of the International Petroleum Technology Conference. Kuala Lumpur, Malaysia.

Appendix

Write your appendix here...