Risk Management in Intelligent Agents

Xun Wang

A Thesis submitted for the degree of Doctor of Philosophy

Faculty of Engineering and Information Technology

University of Technology, Sydney

2012

Certificate of Authorship and Originality

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Student

Acknowledgements

Working on my doctoral research has been one of the most memorable experiences of my life. I would like to thank all people who have helped and inspired me during the past five years of my study.

First and foremost, I would like to thank my supervisor, Professor Mary-Anne Williams. Without your inspiration and encouragement, I would not even contemplate working on my Ph.D research project. Throughout my Ph.D career, you have shown me what would be possible and steer me towards my research goals. You have given me great support in developing my main research ideas; challenged me on many of my immature thoughts. You have pushed me to reach a level of scientific standards that I would not be able to achieve alone. In addition to my main research project, you have also encouraged me to explore new technologies and supported me on fun and challenging projects such as RoboCup. These "side" projects have kept me sane when my main research topic overwhelmed me at times. "Research should be fun" is what I have learned from you. You are a great mentor and a friend. I am deeply grateful for what you have done for me.

Second, I would like to thank many of my colleagues for reviewing my thesis and collaborating with me on many projects. Birger Johansson was the first person who read my thesis draft while it was still an incoherent piece of text. Your suggestions were very useful. I enjoyed many insightful discussions with you while you were visiting our laboratory. I also had the great opportunity working with Chris Stanton on the RoboCup project which I enjoyed enormously. Thanks for correcting many of my grammatical errors in my thesis. Your comments on how to structure my work has greatly helped me. Sajjad Haider taught me a great deal on machine learning and Bayesian inference techniques. Your inputs on my thesis have further improved some of my research ideas. I would also like to thank Benjamin Johnston for his great effort going through my final thesis draft in details. Many of your comments were invaluable. I am also in debt to Shan Chen for keeping my spirit up when I was exhausted at times. I also had pleasure working with Jebrin Al-Sharawneh. Our discussions on what is risk were challenging and insightful.

Third, I would like to thank Thierry Denoeux for our discussion on Belief Functions and Transferable Belief Model during KSEM 2010 conference and many anonymous reviewers for their comments and suggestions on the ideas presented in my research papers. My thesis examiners Didier Dubois and Pavlos Peppas have given me valuable critiques and insightful suggestions. I am forever grateful for your helps! I would also like to thank IBM for their financial support during the fourth year of my Ph.D career.

Finally, I would like to thank my parents, my wife Lily and my daughter Alison for putting up with me for such a long time. Without your support and understanding, this thesis would not be possible.

Table of Contents

List of Figures

List of Tables

Abstract

This thesis presents the development of a generalised risk analysis, modelling and management framework for intelligent agents based on the state-of-art techniques from knowledge representation and uncertainty management in the field of Artificial Intelligence (AI). Assessment and management of risk are well established common practices in human society. However, formal recognition and treatment of risk are not usually considered in the design and implementation of (most existing) intelligent agents and information systems. This thesis aims to fill this gap and improve the overall performance of an intelligent agent. By providing a formal framework that can be easily implemented in practice, my work enables an agent to assess and manage relevant domain risks in a consistent, systematic and intelligent manner.

In this thesis, I canvas a wide range of theories and techniques in AI research that deal with uncertainty representation and management. I formulated a generalised concept of risk for intelligent agents and developed formal qualitative and quantitative representations of risk based on the Possible Worlds paradigm. By adapting a selection of mature knowledge modelling and reasoning techniques, I develop a qualitative and a quantitative approach of modelling domains for risk assessment and management. Both approaches are developed under the same theoretical assumptions and use the same domain analysis procedure; both share a similar iterative process to maintain and improve domain knowledge base continuously over time. Most importantly, the knowledge modelling and reasoning techniques used in both approaches share the same underlying paradigm of Possible Worlds. The close connection between the two risk modelling and reasoning approaches leads us to combine them into a hybrid, multi-level, iterative risk modelling and management framework for intelligent agents, or HiRMA, that is generalised for risk modelling and management in many disparate problem domains and environments. Finally, I provide a top-level guide on how HiRMA can be implemented in a practical domain and a software architecture for such an implementation. My work lays a solid foundation for building better decision support tools (with respect to risk management) that can be integrated into existing or future intelligent agents.