Smart Innovation, Systems and Technologies 96

Gordan Jezic
Yun-Heh Jessica Chen-Burger
Robert J. Howlett
Lakhmi C. Jain · Ljubo Vlacic
Roman Šperka *Editors*

Agents and Multi-Agent Systems: Technologies and Applications 2018

Proceedings of the 12th International Conference on Agents and Multi-Agent Systems: Technologies and Applications (KES-AMSTA-18)

KES International

Springer

Smart Innovation, Systems and Technologies

Volume 96

Series editors

The Smart Innovation, Systems and Technologies book series encompasses the topics of knowledge, intelligence, innovation and sustainability. The aim of the series is to make available a platform for the publication of books on all aspects of single and multi-disciplinary research on these themes in order to make the latest results available in a readily-accessible form. Volumes on interdisciplinary research combining two or more of these areas is particularly sought.

The series covers systems and paradigms that employ knowledge and intelligence in a broad sense. Its scope is systems having embedded knowledge and intelligence, which may be applied to the solution of world problems in industry, the environment and the community. It also focusses on the knowledge-transfer methodologies and innovation strategies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of disciplines from science, technology, business and the humanities. The series will include conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

More information about this series at http://www.springer.com/series/8767

Gordan Jezic · Yun-Heh Jessica Chen-Burger
Robert J. Howlett · Lakhmi C. Jain
Ljubo Vlacic · Roman Šperka
Editors

Agents and Multi-Agent Systems: Technologies and Applications 2018

Proceedings of the 12th International Conference on Agents and Multi-Agent Systems: Technologies and Applications (KES-AMSTA-18)

Springer

Editors
Gordan Jezic
University of Zagreb, Faculty of Electrical Engineering and Computing
Zagreb, Croatia

Yun-Heh Jessica Chen-Burger
The Heriot-Watt University
Edinburgh
Scotland, UK

Robert J. Howlett
Bournemouth University
Poole, UK

and

KES International
Shoreham-by-Sea, UK

Lakhmi C. Jain
Centre for Artificial Intelligence, Faculty of Engineering and Information Technology
University of Technology Sydney
Sydney, NSW, Australia

and

Faculty of Science, Technology and Mathematics
University of Canberra
Canberra, ACT, Australia

and

KES International
Shoreham-by-Sea, UK

Ljubo Vlacic
Griffith Sciences - Centres and Institutes
Griffith University
South Brisbane, QLD, Australia

Roman Šperka
Department of Business Economics and Management and Silesian University in Opava, School of Business Administration in Karvina
Karvina, Czech Republic

ISSN 2190-3018 ISSN 2190-3026 (electronic)
Smart Innovation, Systems and Technologies
ISBN 978-3-319-92030-6 ISBN 978-3-319-92031-3 (eBook)
https://doi.org/10.1007/978-3-319-92031-3

Library of Congress Control Number: 2018944389

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume contains the proceedings of the 12th KES Conference on Agent and Multi-Agent Systems: Technologies and Applications (KES-AMSTA 2018) which will be held in Gold Coast, Australia, between 20 and 22 June 2018. The conference was organized by KES International, its focus group on agent and multi-agent systems and University of Zagreb, Faculty of Electrical Engineering and Computing. The KES-AMSTA conference is a subseries of the KES conference series.

Following the success of previous KES conferences on Agent and Multi-Agent Systems: Technologies and Applications, held in Vilamoura, Portugal (KES-AMSTA 2017), Puerto de la Cruz, Tenerife, Spain (KES-AMSTA 2016), Sorrento, Italy (KES-AMSTA 2015), Chania, Greece (KES-AMSTA 2014), Hue, Vietnam (KES-AMSTA 2013), Dubrovnik, Croatia (KES-AMSTA 2012), Manchester, UK (KES-AMSTA 2011), Gdynia, Poland (KES-AMSTA 2010), Uppsala, Sweden (KES-AMSTA 2009), Incheon, Korea (KES-AMSTA 2008) and Wroclaw, Poland (KES-AMSTA 2007), the conference featured the usual keynote talks, oral presentations and invited sessions closely aligned to its established themes.

KES-AMSTA is an international scientific conference for discussing and publishing innovative research in the field of agent and multi-agent systems and technologies applicable in the digital and knowledge economy. The aim of the conference is to provide an internationally respected forum for both the research and industrial communities on their latest work on innovative technologies and applications that is potentially disruptive to industries. Current topics of research in the field include technologies in the area of mobile and cloud computing, big data analysis, Internet of Things (IoT), business intelligence, artificial intelligence, social systems, computer embedded systems and nature-inspired manufacturing. Special attention is paid on the feature topics: agent interaction and collaboration, modelling and simulation agents, social networks, business informatics, intelligent agents and multi-agent systems.

The conference attracted a substantial number of researchers and practitioners from all over the world who submitted their papers for main track covering the methodologies of agent and multi-agent systems applicable in the digital and knowledge economy, and three invited sessions on specific topics within the field. Submissions came from 15 countries. Each paper was peer-reviewed by at least two members of the International Programme Committee and International Reviewer Board. Thirty-four papers were selected for oral presentation and publication in the volume of the KES-AMSTA 2018 proceedings.

The Programme Committee defined the following main tracks: intelligent agent interaction and collaboration, modelling, simulation and mobile agents, and agent communication and social networks. In addition to the main tracks of the conference, there were the following invited sessions: design and implementation of intelligent agents and multi-agent systems, business informatics and business process management.

Accepted and presented papers highlight new trends and challenges in agent and multi-agent research. We hope that these results will be of value to the research community working in the fields of artificial intelligence, collective computational intelligence, health, robotics, dialogue systems and, in particular, agent and multi-agent systems, technologies, tools and applications.

The Chairs' special thanks go to the following special session organizers: Prof. Lenin G. Lemus-Zúñiga, Universitat Politècnica de València, España; Prof. Arnulfo Alanis Garza, Instituto Tecnológico de Tijuana, México; Prof. Setsuya Kurahashi, University of Tsukuba, Tokyo; Prof. Takao Terano, Tokyo Institute of Technology, Japan; and Prof. Hiroshi Takahashi, Keio University, Japan, for their excellent work.

Thanks are due to the Programme Co-chairs, all Programme and Reviewer Committee members and all the additional reviewers for their valuable efforts in the review process, which helped us to guarantee the highest quality of selected papers for the conference.

We cordially thank all authors for their valuable contributions and all of the other participants in this conference. The conference would not be possible without their support.

April 2018

Gordan Jezic
Jessica Chen-Burger
Robert J. Howlett
Lakhmi C. Jain
Ljubo Vlacic
Roman Šperka

KES-AMSTA-2018 Conference Organization

KES-AMSTA 2018 was organized by KES International—Innovation in Knowledge-Based and Intelligent Engineering Systems.

Honorary Chairs

L. Vlacic	Griffith University, Gold Coast, Australia
I. Lovrek	University of Zagreb, Croatia
L. C. Jain	University of South Australia, Adelaide

Conference Co-chairs

G. Jezic	University of Zagreb, Croatia
J. Chen-Burger	Heriot-Watt University, Scotland, UK

Executive Chair

R. J. Howlett	Bournemouth University, UK

Programme Co-chairs

M. Kusek	University of Zagreb, Croatia
R. Sperka	Silesian University in Opava, Czech Republic

Publicity Chair

P. Skocir	University of Zagreb, Croatia

International Programme Committee

Koichi Asakura	Daido University, Japan
Ahmad Taher Azar	Faculty of Computers and Information, Benha University, Egypt
Marina Bagić Babac	University of Zagreb, Croatia
Dariusz Barbucha	Gdynia Maritime University, Poland
Grażyna Brzykcy	Poznań University of Technology, Department of Control and Information Engineering, Poland
Frantisek Capkovic	Slovak Academy of Sciences, Bratislava, Slovakia
Matteo Cristani	Universita di Verona, Italy
Ireneusz Czarnowski	Gdynia Maritime University, Poland
Paulina Golinska-Dawson	Poznan University of Technology, Poland
Arnulfo Alanis Garza	Instituto Tecnológico de Tijuana. México
Mirjana Ivanovic	University of Novi Sad, Serbia
Dennis Jarvis	Central Queensland University, Australia
Piotr Jedrzejowicz	Gdynia Maritime University, Poland
Dragan Jevtic	University of Zagreb, Croatia
Vicente Julian	Universitat Politecnica de Valencia, Spain
Arkadiusz Kawa	Poznan University of Economics and Business, Poland
Adrianna Kozierkiewicz-Hetmańska	Wroclaw University of Science and Technology, Poland
Konrad Kułakowski	AGH University of Science and Technology, Krakow, Poland
Setsuya Kurahashi	University of Tsukuba, Tokyo
Mario Kusek	University of Zagreb, Croatia
Kazuhiro Kuwabara	Ritsumeikan University, Japan
Jooyoung Lee	Innopolis University, Russia
Marin Lujak	IMT Lille Douai, Douai, France
Evgeni Magid	Kazan Federal University, Russia
Manuel Mazzara	Innopolis University, Russia
Daniel Moldt	University of Hamburg, Germany
Ngoc Thanh Nguyen	Wroclaw University of Technology, Poland
Vedran Podobnik	University of Zagreb, Croatia
Radu-Emil Precup	Politehnica University of Timisoara, Romania
Nafees Qamar	Southwest University, China

Victor Rivera	Innopolis University, Russia
Ewa Ratajczak-Ropel	Gdynia Maritime University, Poland
Katka Slaninova	School of Business Administration in Karvina, Silesian University in Opava, Czech Republic
Silvia Rossi	University of Naples Federico II, Italy
Roman Šperka	Silesian University in Opava, Czech Republic
Darko Stipaničev	University of Split, Croatia
Ryszard Tadeusiewicz	AGH University of Science and Technology, Krakow, Poland
Hiroshi Takahashi	Keio University, Japan
Yasufumi Takama	Tokyo Metropolitan University, Japan
Takao Terano	Tokyo Institute of Technology, Japan
Krunoslav Tržec	Ericsson Nikola Tesla, Croatia
Taketoshi Ushiama	Kyushu University, Japan
Jordi Vallverdú	Universitat Autonoma de Barcelona, Spain
Toyohide Watanabe	Nagoya Industrial Science Research Institute, Japan
Izabela Wierzbowska	Gdynia Maritime University, Poland
Mahdi Zargayouna	University of Paris-Est, IFSTTAR, France
Lenin G. Lemus-Zúñiga	Universitat Politècnica de València. España

Invited Session Chairs

Business Process Management

Roman Šperka	Silesian University in Opava, Czech Republic

Agent-Based Modelling and Simulation

Roman Šperka	Silesian University in Opava, Czech Republic

Business Informatics

Setsuya Kurahashi	University of Tsukuba, Japan
Takao Terano	Tokyo Institute of Technology, Japan
Hiroshi Takahashi	Keio University, Japan

Anthropic-Oriented Computing

Manuel Mazzara	Innopolis University, Russia
Jooyoung Lee	Innopolis University, Russia
Victor Rivera	Innopolis University, Russia

The design and Implementation of Intelligent Agents and Multi-agent Systems

Lenin G. Lemus-Zuniga	Universitat Politecnica de Valencia, Spain
Arnulfo Alanis Garza	Instituto Tecnologico de Tijuana, Mexico

Contents

Intelligent Agent Interaction and Collaboration

Human-Agent Collaboration: A Goal-Based BDI Approach 3
Salma Noorunnisa, Dennis Jarvis, Jacqueline Jarvis, and Marcus Watson

Evolution Direction of Reward Appraisal in Reinforcement Learning Agents 13
Masaya Miyawaki, Koichi Moriyama, Atsuko Mutoh, Tohgoroh Matsui, and Nobuhiro Inuzuka

A General Framework for Formulating Adjustable Autonomy of Multi-agent Systems by Fuzzy Logic 23
Salama A. Mostafa, Rozanawati Darman, Shihab Hamad Khaleefah, Aida Mustapha, Noryusliza Abdullah, and Hanayanti Hafit

Agent-Based System for Context-Aware Human-Computer Interaction 34
Renato Soic, Pavle Skocir, and Gordan Jezic

Agent-Oriented Smart Factory (AOSF): An MAS Based Framework for SMEs Under Industry 4.0 44
Fareed Ud Din, Frans Henskens, David Paul, and Mark Wallis

Modeling, Simulation and Mobile Agents

Agent-Based Approach for Energy-Efficient IoT Services Discovery and Management 57
Petar Krivic, Pavle Skocir, and Mario Kusek

Agent-Based Modeling and Simulation for Two-Dimensional Spatial Competition 67
Masashi Miura and Hidetoshi Shiroishi

Agent Based Simulation of Network Routing: Reinforcement Learning Comparison 76
Krešimir Čunko, Marin Vuković, and Dragan Jevtić

Dispatching Strategies for Dynamic Vehicle Routing Problems 87
Besma Zeddini and Mahdi Zargayouna

Securing Mobile Agents, Stationary Agents and Places in Mobile Agents Systems 97
Donies Samet, Farah Barika Ktata, and Khaled Ghedira

How Research Achievements Can Influence Delivering of a Course - Siebog Agent Middleware 110
Milan Vidaković, Mirjana Ivanović, Dejan Stantić, and Jovana Vidaković

Agent Communication and Social Networks

Sending Messages in Social Networks 123
Matteo Cristani, Francesco Olivieri, Claudio Tomazzoli, and Guido Governatori

ER-Agent Communication Languages and Protocol for Large-Scale Emergency Responses 134
Mohd Khairul Azmi Hassan and Yun-Heh Chen-Burger

Towards a Logical Framework for Diagnostic Reasoning 144
Matteo Cristani, Francesco Olivieri, Claudio Tomazzoli, and Margherita Zorzi

Scalability of Dynamic Lighting Control Systems 156
Leszek Kotulski and Igor Wojnicki

Automatic Detection of Device Types by Consumption Curve 164
Claudio Tomazzoli, Matteo Cristani, Simone Scannapieco, and Francesco Olivieri

Business Process Management

Advantages of Application of Process Mining and Agent-Based Systems in Business Domain 177
Michal Halaška and Roman Šperka

Modelling the Validation Process of Enterprise Software Systems 187
Robert Bucki and Petr Suchánek

Design and Implementation of Intelligent Agents and Multi-Agent Systems I

Multi-agent System for Forecasting Based on Modified Algorithms of Swarm Intelligence and Immune Network Modeling 199
Galina A. Samigulina and Zhazira A. Massimkanova

Multi-Agent System Model for Diagnosis of Personality Types 209
Margarita Ramírez Ramírez, Hilda Beatriz Ramírez Moreno, Esperanza Manrique Rojas, Carlos Hurtado, and Sergio Octavio Vázquez Núñez

Towards a Multi-Agent System for an Informative Healthcare Mobile Application 215
Carlos Hurtado, Margarita Ramirez Ramirez, Arnulfo Alanis, Sergio Octavio Vazquez, Beatriz Ramirez, and Esperanza Manrique

Proposal of a Bootcamp's User Activity Dashboard Based on MAS 220
Lenin G. Lemus-Zúñiga, Valeria Alexandra Haro Valle, José-V. Benlloch-Dualde, Edgar Lorenzo-Sáez, Miguel A. Mateo Pla, and Jorge Maldonado-Mahauad

Toward to an Electric Monitoring Platform Based on Agents 231
Jorge E. Luzuriaga, Guillermo Cortina Rodríguez, Karolína Janošová, Monika Borova, Miguel Ángel Mateo Pla, and Lenin-G. Lemus-Zúñiga

Design and Implementation of Intelligent Agents and Multi-Agent Systems II

A Cooperative Agent-Based Management Tool Proposal to Quantify GHG Emissions at Local Level 243
Edgar Lorenzo-Sáez, José-Vicente Oliver-Villanueva, Jorge E. Luzuriaga, Miguel Ángel Mateo Pla, Javier F. Urchueguía, and Lenin-G. Lemus-Zúñiga

A Proposal to Improve the Usability of Applications for Users with Autism Considering Emotional Aspects 253
Ángeles Quezada, Reyes Juarez-Ramirez, Arnulfo Alanís Garza, Bogart Yail, Sergio Magdaleno, and Eugenia Bermudez

Towards a Model Based on Agents for the Detection of Behavior Patterns in Older Adults Who Start Using ICT 261
Consuelo Salgado Soto, Maricela Sevilla Caro, Ricardo Rosales Cisneros, Margarita Ramírez Ramírez, Hilda Beatriz Ramírez Moreno, and Esperanza Manrique Rojas

Intelligent Agents as Support in the Process of Disease Prevention Through Health Records 269
Hilda Beatriz Ramirez Moreno, Margarita Ramírez Ramírez, Esperanza Manrique Rojas, Nora del Carmen Osuna Millán, and Maricela Sevilla Caro

Agent-Based Model as a Provider of Medical Services in Tijuana Mexico 275
Ricardo Rosales, Nora Osuna-Millan, Consuelo Salgado-Soto, Carlos Flores-Sanchez, Juan Meza-Fregoso, and Arnulfo Alanis

Business Informatics

Understanding the Potential Value of Digitization for Business – Quantitative Research Results of European Experts 287
Christopher Reichstein, Ralf-Christian Härting, and Pascal Neumaier

A Method of Knowledge Extraction for Response to Rapid Technological Change with Link Mining 299
Masashi Shibata and Masakazu Takahashi

Agent-Based Gaming Approach for Electricity Markets 311
Setsuya Kurahashi

Finding the Better Solutions for the Smart Meter Gateway Placement in a Power Distribution System Through an Evolutionary Algorithm 321
Ryoma Aoki and Takao Terano

Japanese Health Food Market Trend Analysis 331
Yoko Ishino

Simulation of the Effect of Financial Regulation on the Stability of Financial Systems and Financial Institution Behavior 341
Takamasa Kikuchi, Masaaki Kunigami, Takashi Yamada, Hiroshi Takahashi, and Takao Terano

Author Index 355

Intelligent Agent Interaction and Collaboration

Human-Agent Collaboration: A Goal-Based BDI Approach

Salma Noorunnisa[1], Dennis Jarvis[1(✉)], Jacqueline Jarvis[1], and Marcus Watson[2]

[1] Centre for Intelligent Systems, Central Queensland University, 160 Ann Street, Brisbane 4000, Australia
{s.noorunnisa,d.jarvis,j.jarvis}@cqu.edu.au
[2] School of Psychology, University of Queensland, St. Lucia 4067, Australia
m.watson2@uq.edu.au

Abstract. The Belief-Desire-Intention (BDI) model of agency has been a popular choice for the modelling of goal-based behaviour for both individual agents and more recently, teams of agents. Numerous frameworks have been developed since the model was first proposed in the early 1980s. However, while the more recent frameworks support a delegative model of agent/agent and human/agent collaboration, no frameworks support a general model of collaboration. Given the importance of collaboration in the development of practical semi-autonomous agent applications, we consider this to constitute a major limitation of traditional BDI frameworks. In this paper, we present GORITE, a novel BDI framework that by employing explicit goal representations, overcomes many of the limitations of traditional frameworks. In terms of human/agent collaboration, key requirements are identified and through the use of a representative but simple example, the ability of GORITE to address those requirements is demonstrated.

Keywords: Human-agent collaboration · BDI · Multi-agent systems

1 Introduction

At a recent workshop on Human-Autonomy Teaming (HAT), having a shared mental model was identified as being essential if HAT systems are to deliver the levels of trust and explanatory capability required for military operations [1]. Furthermore, the Situation Awareness-Based Agent Transparency (SAT) Model developed by Chen et al. [2] was identified as providing a suitable conceptual framework for future HAT research. SAT visualisation agents [2] are able to provide operators with situation awareness of an evolving mission environment. This support is at three levels:

(1) What's going on and what is the agent trying to achieve?
(2) Why does the agent do it?
(3) What should the operator expect to happen?

In addressing these questions, a SAT agent draws on its desires and intentions at Level 1 and its beliefs at Level 2. However, while the SAT model is inspired by the

G. Jezic et al. (Eds.): KES-AMSTA-18 2018, SIST 96, pp. 3–12, 2019.
https://doi.org/10.1007/978-3-319-92031-3_1

Belief-Desire-Intention (BDI) model of agency, the creation of SAT agents grounded in the BDI model of agency is a research issue yet to be explored. Rather, the focus of Chen's work has been the visualisation of information pertaining to the SAT levels and the demonstration through controlled experimentation that operator performance and trust in automation is enhanced through such visualisation. Chen has proposed that the research community should continue with that agenda and we are in agreement. However, we see a major opportunity for a complementary research program with a focus of developing a BDI software framework that explicitly supports human/agent collaboration through the use of the SAT model. As explained below, this will require a framework that provides explicit representation of beliefs, desires and intentions in order to enable the agent to reflect on and explain its actions and to enable humans to dynamically modify agent behaviour.

This represents a significant departure from traditional BDI agent frameworks and we propose to use GORITE, a novel open-source BDI framework developed by Rönnquist [3] for this purpose. GORITE itself is a mature, open source and fully functional software framework, as evidenced by the case studies presented in [3]. In particular, the manufacturing control case study is a reimplementation of an earlier commercial application developed using the JACK Teams BDI framework. However, the version of GORITE and the case studies in [3] support and focus on autonomous BDI behavior. For effective human/BDI agent collaboration, extensions to both the BDI model and to the GORITE framework are required. Our intent is to tackle this iteratively, with each iteration involving model extension, framework realization and application development. The application (waypoint traversal) presented in this paper represents the key functionality for our domain of interest – war gaming using semi-automated computer generated forces. As such, it provides an ideal example for demonstrating the effectiveness of the extensions to both the BDI model and the GORITE framework.

In the remainder of this paper, we will first discuss the BDI model, its limitations with respect to human/agent collaboration and how these can be overcome with GORITE. We then demonstrate how effective human/agent collaboration can be achieved for SAT Levels 1 and 2 using the GORITE framework. In this regard, a simple but representative example (waypoint traversal by a platoon) is employed, in order to maintain focus on the key collaboration requirements. SAT Level 3 functionality and agent/human collaboration are out of scope for the current research activity. Note that the innovation in this work lies in its extension of the BDI Model of agency to accommodate human/agent collaboration and the realization of this extended model in the GORITE BDI framework.

2 The BDI Model

The BDI model is concerned with how an agent makes rational decisions about the actions that it performs through the employment of

(1) Beliefs about their environment, other agents and themselves,
(2) Desires that they wish to satisfy and
(3) Intentions to act towards the fulfilment of selected desires.

The model has its origin in Bratman's theory of human practical reasoning [4]. Bratman's ideas were first formalised by Rao and Georgeff [5] who subsequently proposed an abstract architecture in which beliefs, desires and intentions were explicitly represented as global data structures and where agent behaviour is event driven. However, while this conceptualisation faithfully captured Bratman's theory, it did not constitute a practical system for rational reasoning. In order to ensure computational tractability, they proposed the following representational changes [6]:

- Only beliefs about the current state of the world are represented explicitly
- Information about the means of achieving certain future world states (desires) and the options available to an agent are represented as plans.
- A particular desire may be realizable by multiple plans but an agent must select one plan to pursue.
- Plans either succeed or fail; if a plan fails, then the desire which is being pursued may be reconsidered.
- Intentions are represented implicitly by the collection of currently active plans
- Desires are referred to as goals, which are represented as events. Goals have only a transient representation, acting as triggers for plan invocations.

These considerations led to the following execution model:

```
repeat
  wait for the next goal event;
  select (on the basis of current beliefs) a plan to
  achieve the current goal;
  execute the selected plan;
  update beliefs;
end repeat
```

This execution model has provided the conceptual basis for all major research and commercial BDI implementations, in particular PRS, dMARS and JACK [3].

In the traditional BDI execution model outlined above, plans consist of steps that are specified using a framework dependent plan language; these steps may involve the posting of further goal events (or the reposting of the current goal event). More than one plan may be applicable for the achievement of a particular goal – this set of plans is called the applicable set. The selection of a plan to execute from the applicable set is based on the currently held beliefs of the agent and may involve explicit (meta-level) reasoning.

Since its inception, the Belief-Desire-Intention (BDI) model of agency has underpinned many successful agent applications and has been identified as one of the preferred vehicles for the delivery of industry strength, knowledge rich, intelligent agent applications [7]. As originally conceived by Bratman, the model was intended as a means to determine how an agent should act in a situated environment. The early applications of the model reflect this focus on situated, autonomous behaviour, but within constrained technical domains, e.g. space shuttle fault diagnosis [8]. While this has continued to be a focal point for applications, as evidenced by its use in manufacturing system and UAV control, commercial success has been achieved in its application to human behaviour modelling in war gaming, where credible entity behaviour is an essential requirement for an effective military game [3]. The wargaming examples have necessitated significant extensions to the BDI execution model as originally formulated by Rao and Georgeff. In particular, the model has been extended in JACK Teams [9] to accommodate teams (such as platoons and companies) as distinct entities with their own beliefs, desires and intentions and in CoJACK [9] to provide agents with an explicit cognitive architecture to ground reasoning. However, these extensions retain a key feature of the Rao and Georgeff model, namely that the goals are not represented as explicit, persistent entities, but rather as transitory events. This makes reasoning about an agent's intentions – for example, whether to continue or discontinue with the current goal or how a goal should be resourced – problematic.

While the extensions to the BDI execution model embodied in both JACK Teams and CoJACK have significantly extended the range of problems that can be effectively addressed by BDI agents, these problems remain characterized by a requirement for autonomous execution potentially supported by delegation. The applicability of the BDI execution model becomes problematic when human/agent collaboration is required. If the collaboration involves only simple delegation, with execution being managed by either the human or the agent, then the delegation model supported by JACK Teams will suffice. However a more comprehensive collaboration model is required if goal/belief inspection/management is required of the collaboration. For example, the provision of such functionality would significantly increase the amount of behaviour that could be delegated to agents (and teams of agent) in theatre-level wargames, as the puckster would have the ability to dynamically interact with the agents.

To summarise, in terms of human/agent collaboration, the BDI execution model exhibits the following limitations:

- Interruption of plans is not supported.
- Goal representation is implicit and transient, with goals modelled as events that are not persisted. Consequently, goals are not inspectable.
- Depending on how beliefs are stored, they may be inspectable. However, no distinction is made in the BDI model between individual agent beliefs, shared agent beliefs and beliefs that are shared by agents that are collaborating on a particular goal execution.

Additional insight can be gained into the requirements of human/agent collaboration by considering the more general problem of Activity Based Computing (ABC), where human/human collaboration is mediated by a shared computational workspace. Furthermore, one can reasonably expect the key requirements for human/agent collaboration to be a subset of the requirements for ABC. Activity Based Computing was conceived by Norman, one of the pioneers of HCI. However, realisation of the concept was left to others, most notably Bardram and his colleagues [10]. ABC is of particular relevance to the current research problem because it is concerned with the support that people need when working on a shared computational activity. While our interest is in human/agent collaboration rather than human/human collaboration, the key requirements that ABC imposes on shared activities, namely suspension and resumption, context awareness and inspectability also apply to goal executions that are to be shared with humans. It is also of interest that Norman [11] has identified goals as being an appropriate conceptualization for reasoning about activity. However, this conceptualization is not present in Bardram's current work.

Based on Chen's SAT model, Bardram's work and reflection on our experience in developing multi-agent applications for military wargaming, the following key functional requirements for effective human/agent collaboration have been identified. In particular, in order to satisfy SAT Level 1 and 2 functionality, a human must be able to

(1) Delegate goal execution to an agent
(2) Suspend and resume a particular goal execution
(3) Determine why an agent has chosen to pursue a particular course of action.
(4) Inspect beliefs relevant to a particular goal execution and if appropriate, make modifications.
(5) Inspect the goals that an agent has committed to pursue and if necessary, add new goals, delete existing goals or modify the execution order

Additional functionality such as goal replay, goal re-execution with modified context and persistence of goal execution state may be beneficial in some circumstances and particularly at SAT Level 3. However, as our immediate focus is SAT Levels 1 and 2, such functionality is deemed to be out of scope. Also note that requirement 2 (suspension and resumption of goal execution) is a prerequisite for requirements 3-5 and hence constitutes the key focus of this paper.

3 Gorite

GORITE is an open source Java framework that provides class level support for the development of agent applications that involve teams of BDI agents. Agents in GORITE are modelled as Java classes that extend the structural framework classes (Performer and Team). Agent behaviour is specified in terms of goal-based process models, which are code-level constructs that employ the behavioural framework classes (Goal and its sub-classes). Below is the specification from the case study for the platoon performer's path traversal goal.

```
private Goal traversePath() {
    return new SequenceGoal(TRAVERSE_PATH, new Goal[]{
        new Goal("process percept") {
            @Override
            public Goal.States execute(Data d) {
                System.err.println("Execution started");
                Path p = (Path) d.getValue(PERCEPT);
                d.setValue(PATH, p);
                String ename = (String) d.getValue(EXECUTION);
                Execution e = etable.get(ename);
                e.state = State.RUNNING;
                String m = timeStamp()+ename+" : Execution Started";
                etable.inform( ename, m, e.state.name() );
                etable.inform( "log", m, null );
                return Goal.States.PASSED;
            }
        },
        new LoopGoal("visit waypoints", new Goal[]{
            traverseSegment(),
            trackProgress(),
            checkpoint(),
        })
    });
}
```

Note that the path traversal goal specifies both the activities required to achieve the goal and the coordination requirements for those activities. Both facets are specified explicitly and uniformly using GORITE goal class instances (e.g. `Goal`, `SequenceGoal`, `LoopGoal` in the method above). The resulting traversal goal instance can then be executed on behalf of the goal owner by a separate executor object, which traverses the goal instance graph and at each node (which is an object of type `Goal`) invoking its `execute()` method. BDI execution semantics are preserved, with the agent still able to choose between courses of action to achieve a goal or to reconsider how a goal might be achieved. Team goals are specified in terms of roles which are filled by team members; team goal execution is then managed by a single executor on behalf of all the participating team members. During the goal graph traversal process, the executor makes available to the participants in the execution a shared data context, thus providing for a clear separation between an agent's individual beliefs and those that it shares with other agents involved in the goal execution.

In traditional BDI frameworks, execution is agent focused – an individual agent (or an agent team) determines what plan to perform in order to achieve its current goal. This determination is done in the absence of any explicit representation of currently active or future intentions. In GORITE, the focus is shifted from an individual agent pursuing its current goal to an individual agent being a participant in the achievement of a larger system-level goal. We would argue that any practical agent framework should provide support for both perspectives. That is, an agent needs to be able to operate as part of a larger whole while at the same time, progressing its own goals if appropriate. In GORITE, this latter behaviour is supported through the concept of a ToDo group. Each performer can maintain a ToDo group, which is a list of the intentions that it is currently pursuing or has decided to pursue in the near future. Within a ToDo group, only one intention is progressed during a time slice – that is the intention that is at the top of the list. However, prior to the executor progressing the top intention at the beginning of a time slice, meta-level reasoning can be invoked to determine which intention is to be progressed in the next time slice. ToDo groups can also be used to model reactive behaviour, including user interaction. In this respect, GORITE provides a `Perceptor` class that can be used by a performer to add goals to its ToDo group when particular events occur. In [3], perceptors were used to model incoming manufacturing orders and requests for sensor team reformation. However, they also provide a convenient mechanism for user-initiated goal execution. The traversal goal is added to the company performer's ToDo group via the `Company.start()` method:

```
public void start(String ename,String gname,Object percept,Data d)
{
     System.err.println("Execution to be started");
     Execution e = new Execution(ename);
     e.request = Request.START;
     etable.put(ename, e);
     String m = timeStamp()+ename+" : Traversal requested";
     etable.inform( ename, m, e.state.name() );
     etable.inform( "log", m, null );
     d.setValue(EXECUTION, ename);
     Perceptor perceptor = perceptors.get(gname);
     perceptor.perceive(percept, d);
 }
```

The percept object is of type `Path` and contains the waypoints that the platoon is to visit. This object is added to the data context (`d`) for the goal execution within the `Perceptor.perceive()` method as a data element with the default name of `PERCEPT`. For convenience, the path object was made available to the goal execution as an element in the data context called `PATH` in the traversal goal definition provided earlier. The `perceive()` method also adds the goal to the company's ToDo group. In this instance, `start()` is invoked by a method chain originating in the action listener for the Start button in the application GUI. Note that multiple goals can be added to the ToDo group and that these goals can be executed either sequentially, or through the use of meta-goals, concurrently. For a more complete description of the GORITE execution model, the reader is referred to [3].

4 Human/Agent Collaboration

In the previous section, it has been demonstrated how goal execution can be initiated by a human using the GORITE preceptor. While this constitutes an essential first step in the achievement of human/agent collaboration, it constitutes only one of the requirements identified in Sect. 2. As indicated in Sect. 2, the key requirement in terms of providing effective human/agent collaboration is requirement 2 – suspension and resumption of goal execution.

In GORITE goal execution is time-sliced and within a ToDo group, meta-goals can be employed to determine which goal in the ToDo group should be progressed next. However, one then has the problem of determining how to suspend goal execution, perhaps by executing a blocking goal and more importantly, when to suspend execution. The ideal is for goal execution to be interrupted at well-defined points which we refer to as checkpoints and this is the approach that we have employed in the waypoint traversal example. As indicated in the traversal goal definition in the previous section, a checkpoint goal is performed whenever a waypoint is reached. This goal passes if there are no user requests. If there is a suspension request, then the goal blocks until a resumption request is issued by the user:

```
Goal checkpoint() {
    return new Goal( "checkpoint" ) {
        @Override
        public Goal.States execute(Data d) {
            String ename = (String) d.getValue(EXECUTION);
            Execution e = etable.get(ename);
            switch (e.request) {
                case PAUSE:
                  if (e.state == State.RUNNING) {
                    System.err.println("Execution paused");
                    e.state = State.PAUSED;
                    String m1 = timeStamp()+ename+" : paused";
                    etable.inform( ename, m1, e.state.name() );
                  }
                  return Goal.States.BLOCKED;
                case CONTINUE:
                  if (e.state == State.PAUSED) {
                    System.err.println("Execution resumed");
                    e.state = State.RUNNING;
                    String m2 = timeStamp()+ename+" : resumed";
                    etable.inform( ename, m2, e.state.name() );
                  }
                  return Goal.States.PASSED;
            }
        }
    };
```

While goal execution is blocked, the user is able to inspect and modify the data context, which contains the data elements relevant to the goal execution. In our example, the relevant data are the waypoints, and a simple GUI is provided to enable the user to modify future waypoints:

The data context is also an appropriate structure in which to hold explanations as to how the agent has arrived at the current execution point. In this regard, providing a goal trace would constitute a good starting point. If the ToDo group contains multiple goals, then in a similar way, the ordering of the goals can be modified, existing goals removed and new goals added. An example of ToDo group manipulation is provided in [3].

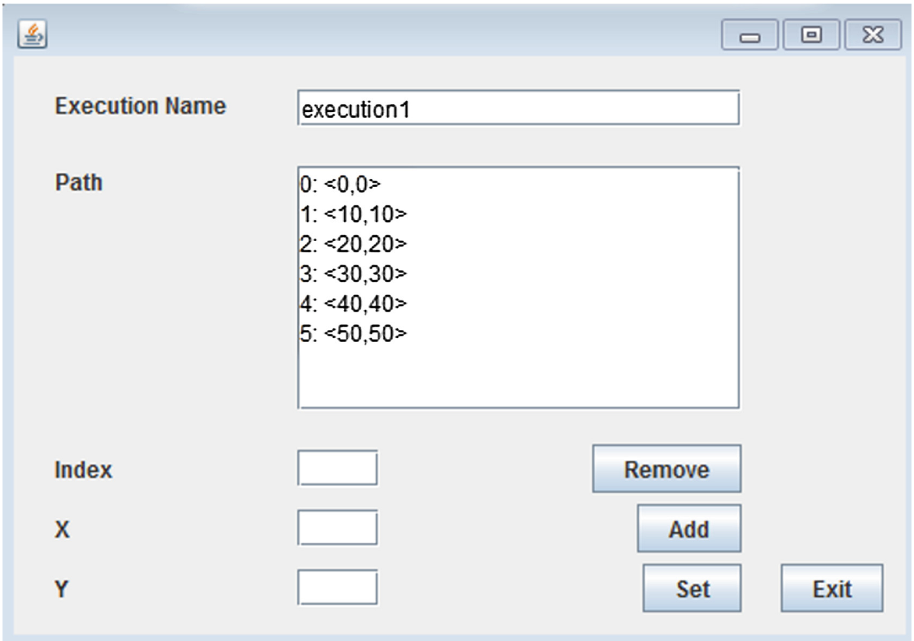

Fig. 1. The waypoint modification GUI

5 Conclusion

The motivation of this work has been to demonstrate that the GORITE BDI framework, through its explicit goal representation and corresponding execution model, supports the key requirements for human/agent collaboration and can be used to develop SAT agents. Through the use of a simple but representative example, the ability for humans to initiate, suspend and resume GORITE agent activity has been demonstrated. The ability to inspect and modify the data associated with the goal execution has been demonstrated. If a GORITE agent is intending to pursue multiple goals (either concurrently or sequentially), the goals in its ToDo group can be manipulated by a collaborating human. This particular aspect of human/agent collaboration was not included in the simplified example used in this paper. From a modelling perspective, the concept of goal execution underpins the requirements identified in Sect. 2 for effective human/agent collaboration. This concept is captured in the application through the `Execution` class and the checkpoint goal, which now form the basis of a generic goal execution capability for the GORITE framework.

Using GORITE to develop effective SAT agents will be an ongoing activity. Chen et al. have demonstrated that the transparency provided by SAT-enabled agents is beneficial in terms of human operator effectiveness. The tasks involved in these studies were relatively straightforward; a key challenge, we believe, will be in the scaling up of human/agent collaboration to address more complex problems. In particular, while GORITE may provide a basic set of building blocks for creating transparent agents, what is not clear is how these agents should be constructed and what additional support should be provided at the framework level. The goal execution concept has proven useful both in this work and in other related applications (manufacturing and medical prescribing) which suggests that such an abstraction is generally useful and should be supported at the framework level. Visualisation is another example where generic support could be provided – for instance using interactive Gantt charts as a vehicle for goal management rather than a conventional GUI-based approach as exemplified by Fig. 1 could be beneficial in terms of the user experience. Also we would see integration with simulation as a key element in the delivery of functionality at SAT Level 3.

References

1. Hill, S.G., Ling, M.F.: Human-Robot Bi-Directional Communication and Human-Autonomy Teaming Workshop Summary. DST-Group-GD-0965, Defence Science and Technology Group, Fishermans Bend (2017)
2. Chen, J.Y.C., Procci, K., Boyce, M., Wright, J., Garcia, A., Barnes, M.: Situation Awareness-Based Agent Transparency. ARL-TR-6905, Army Research Laboratory, Aberdeen Proving Ground, Maryland (2014)
3. Jarvis, D., Jarvis, J., Rönnquist, R., Jain, L.: Multiagent Systems and Applications. Volume 2: Development Using the GORITE BDI Framework. Springer, Heidelberg (2012)
4. Bratman, M.E.: Intention, Plans, and Practical Reason. Harvard University Press, Cambridge (1987)
5. Rao, A.S., Georgeff, M.P.: Modeling rational agents within a BDI architecture. In: Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning. Morgan Kaufman (1991)
6. Rao, A.S., Georgeff, M.P.: BDI agents: from theory to practice. In: Proceedings of the 1st International Conference on Multi-Agent Systems (ICMAS 1995), San Francisco (1995)
7. Jones, R., Wray, R.: Comparative analysis of frameworks for knowledge-intensive intelligent agents. AI Mag. **27**, 57–70 (2006)
8. Georgeff, M.P., Lansky, A.I.: Procedural knowledge. Proc. IEEE **74**, 1383–1398 (1986)
9. AOS Group: AOS Group. http://www.agent-software.com.au
10. Bardram, J.E., Jeuris, S., Houben, S.: Activity-based computing: computational management of activities reflecting human intention. AI Mag. **36**(2), 63–72 (2015)
11. Norman, D.: The Design of Everyday Things: Revised and Expanded Edition. Basic Books, New York (2013)

Evolution Direction of Reward Appraisal in Reinforcement Learning Agents

Masaya Miyawaki[1(✉)], Koichi Moriyama[1], Atsuko Mutoh[1], Tohgoroh Matsui[2], and Nobuhiro Inuzuka[1]

[1] Department of Computer Science, Nagoya Institute of Technology, Nagoya, Japan
m.miyawaki.474@nitech.jp, {moriyama.koichi,atsuko,inuzuka}@nitech.ac.jp
[2] Department of Clinical Engineering, Chubu University, Kasugai, Japan
TohgorohMatsui@tohgoroh.jp

Abstract. Humans appraise the environment in daily life. We are implementing appraisal mechanisms into reinforcement learning agents. One of such mechanisms we proposed is the utility-based Q-learning, which learns behaviors from subjective utilities derived from payoffs the agent gains and a utility-derivation function the agent has. In the previous work, we know that payoff-based evolution brings utility-derivation functions that facilitate mutual cooperation in iterated prisoner's dilemma games. However, the evolution process itself has not yet been known well. In this work, we investigate the process in terms of what determines the evolution direction. We introduce two metrics showing preference of actions based on the evolved subjective utilities, which divide the evolution space into four regions. In each region, the metrics will explain the evolution directions.

Keywords: Multi-agent reinforcement learning · Reward appraisal
Prisoner's dilemma · Genetic algorithm · Evolutionary process

1 Introduction

We humans do not accept the surrounding environment as it is, but appraise it in daily life. Such appraisal mechanisms are changing our behaviors. For example, humans are able to cooperate with each other because we create a kind of rewards for cooperation in our brains [1]. On the other hand, let us consider human-like autonomous computer programs called agents learning their behaviors with reinforcement learning. It is difficult for them to learn cooperative behaviors from given rewards [2] because they are reward-maximizers.

Moriyama [3] proposed the utility-based reinforcement learning concept where an agent learns behaviors from not rewards but *utilities* derived by a utility-derivation function it has. The function is a kind of appraisal mechanism. That work showed a condition of the function giving mutual cooperation in iterated prisoner's dilemma (IPD) games. Later, Moriyama et al. [4] showed

G. Jezic et al. (Eds.): KES-AMSTA-18 2018, SIST 96, pp. 13–22, 2019.
https://doi.org/10.1007/978-3-319-92031-3_2

that such mutually cooperative utility-derivation functions were given by evolutionary computation whose fitness was the sum of rewards. In other words, the cooperative appraisal mechanisms were evolved in an environment where the agents should be reward-maximizers. In addition, interestingly, the evolved functions had a specific structure in the space of utility-derivation functions. However, unfortunately, that work did not yet investigate in detail the evolution process itself showing why and how such a structure evolved.

Therefore, this work investigates the evolution process itself. In order to make it easier, we first introduce two metrics showing preference of actions under an assumption that the opponent takes actions evenly. After that, we investigate the process using the metrics.

Human appraisal mechanisms are fast, intuitive methods for decision making. Hence, for example, similar mechanisms will be needed in robot control in an open environment that requires immediate decisions one after another. This work helps us understand the mechanisms.

This paper consists of six sections. Section 2 is a preliminary section introducing IPD games, Q-learning, and the utility-based Q-learning. In Sect. 3, we propose the metrics. In Sect. 4, we show the experiment where the utility-derivation functions evolved and analyze the result in detail with the metrics. Section 5 refers to some related works. Finally, this paper is concluded in Sect. 6.

2 Preliminaries

2.1 Iterated Prisoner's Dilemma Games

A prisoner's dilemma game [5] is a game where two players simultaneously choose their actions, either cooperation (C) or defection (D), and receive payoffs $r \in \{T, R, P, S\}$, respectively. The payoffs satisfy the following conditions: $T > R > P > S$. The relation between their actions and their payoffs is shown in Table 1. This table shows that if each player pursues individual rationality, mutual defection occurs and both of them receive a payoff P smaller than R, the payoff of mutual cooperation.

An iterated prisoner's dilemma (IPD) game is that the players play a prisoner's dilemma game iteratively. The payoffs satisfy $2R > T + S$. IPD is fascinating researchers for decades and gives us a good example where appraisal should be different from payoffs.

Table 1. Payoff table of a prisoner's dilemma game. The players are given payoffs determined by the combination of their actions, i.e., C and D. The row player receives the left payoffs, while the column receives the right ones.

Row\Column	C	D
C	R, R	S, T
D	T, S	P, P

2.2 Q-Learning

Q-learning [6] is one of the most famous reinforcement learning algorithms. In Q-learning, an agent chooses its action a_t based on an action value $Q(s_t, a_t)$ from available actions in the current state s_t at each time step t. After that, the agent receives a reward r_{t+1} from the environment and the state changes to the next s_{t+1}, and then the action value $Q(s_t, a_t)$ is updated as follows:

$$
\begin{aligned}
Q(s_t, a_t) &\leftarrow Q(s_t, a_t) + \alpha\delta_t, \\
\delta_t &\equiv r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t),
\end{aligned}
$$

where α is a learning rate and γ is a discount rate. The agent will learn the optimal behavior without any explicit instructions, i.e., labels.

2.3 Utility-Based Q-Learning

Utility-based Q-learning [3] is an extension of Q-learning where a *subjective utility* u derived from a reward r from a utility-derivation function $u(r)$ is used as follows:

$$
\begin{aligned}
Q(s_t, a_t) &\leftarrow Q(s_t, a_t) + \alpha\delta_t, \\
\delta_t &\equiv u(r_{t+1}) + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t).
\end{aligned}
\tag{1}
$$

In IPD games, Moriyama et al. [4] evolved coefficients of a cubic utility-derivation function and obtained subjective utilities leading to mutual cooperation. We will examine it in Sect. 4.

3 Action Tendency

Moriyama et al. [4] showed that the evolved subjective utility-derivation functions had a specific structure. However, unfortunately, they did not yet investigate in detail the evolution process itself showing why and how such a structure evolved. Therefore, this work investigates the evolution process itself.

Since the evolution changes the subjective utilities of each agent and the utilities change the agent's behaviors, we focus on the relation between the subjective utilities and behaviors in each agent. We first define two metrics, collectively called "Action Tendency", which shows preference of actions under an assumption that the opponent takes actions evenly. The metrics are called "cooperativeness" and "conformity". We next define "Dominant Action Tendency", which determines which metric gives more influence on an action choice.

3.1 Cooperativeness

Let the agent's action be $X \in \{C, D\}$. We define a function u_{cp} for an action X as follows:

$$
u_{cp}(X) \equiv \begin{cases} \dfrac{u(R) + u(S)}{2} & \text{if } X = C, \\[2ex] \dfrac{u(T) + u(P)}{2} & \text{otherwise, i.e., } X = D. \end{cases}
$$

It shows a subjective utility of each action under an assumption that the opponent chooses one of the actions evenly. We say that the agent is cooperative if $u_{cp}(C) \geq u_{cp}(D)$ and defective otherwise.

Next, we define the cooperativeness metric m_{cp}, which shows the relation between $u_{cp}(C)$ and $u_{cp}(D)$, as follows.

$$m_{cp} \equiv \begin{cases} \dfrac{u_{cp}(C)}{u_{cp}(D)} & \text{if } u_{cp}(D) \neq 0, \\ u_{cp}(C) & \text{otherwise.} \end{cases}$$

We say that the agent is cooperative or defective using the metric as follows.

$$\text{The agent is } \begin{cases} \text{cooperative} & \text{if } \begin{cases} m_{cp} \geq 1 \text{ if } u_{cp}(D) > 0, \\ m_{cp} \geq 0 \text{ if } u_{cp}(D) = 0, \\ m_{cp} \leq 1 \text{ otherwise,} \end{cases} \\ \text{defective} & \text{otherwise.} \end{cases} \tag{2}$$

Note that m_{cp} shows the strength of preference. For example, if $m_{cp} \gg 1$ when $u_{cp}(D) > 0$, the agent is strongly cooperative; if $m_{cp} \ll 1$ when $u_{cp}(D) > 0$, it is strongly defective.

3.2 Conformity

Let us consider a statement I that means the agent's action and the opponent's are identical. We define a function u_{cf} for the statement I and $\neg I$ as follows:

$$u_{cf}(I) \equiv \frac{u(R) + u(P)}{2},$$
$$u_{cf}(\neg I) \equiv \frac{u(T) + u(S)}{2}.$$

It shows a preference for taking a same action with its opponent under an assumption that the action pairs happen evenly. We say that the agent is conforming if $u_{cf}(I) \geq u_{cf}(\neg I)$ and anticonforming otherwise.

Next, we define the conformity metric m_{cf}, which shows the relation between $u_{cf}(I)$ and $u_{cf}(\neg I)$, as follows.

$$m_{cf} \equiv \begin{cases} \dfrac{u_{cf}(I)}{u_{cf}(\neg I)} & \text{if } u_{cf}(\neg I) \neq 0, \\ u_{cf}(I) & \text{otherwise.} \end{cases}$$

We say that the agent is conforming or anticonforming using the metric as follows.

$$\text{The agent is } \begin{cases} \text{conforming} & \text{if } \begin{cases} m_{cf} \geq 1 \text{ if } u_{cf}(\neg I) > 0, \\ m_{cf} \geq 0 \text{ if } u_{cf}(\neg I) = 0, \\ m_{cf} \leq 1 \text{ otherwise,} \end{cases} \\ \text{anticonforming} & \text{otherwise.} \end{cases} \tag{3}$$