Honesty, Trust, and Rational Communication in Multiagent Semi-Competitive Environments

LAM Ka Man

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Philosophy
in
Computer Science and Engineering

July, 2004

Abstract

In multiagent environments, agents need to choose their actions. To achieve this, Gmytrasiewicz and Durfee propose the Recursive Modeling Method (RMM). This thesis improves the probabilistic approximation approach in the original design with Recursive Formulas, which enables agents to predict other agents' actions and choose their own actions more accurately.

In multiagent semi-competitive environments, agents have another type of actions that they need to choose. In such environments, competitions and cooperations can both exist. As agents compete with one another, they have incentives to lie. Sometimes, however agents can increase their utilities by cooperating with each other, then they have incentives to tell the truth. Therefore, being a receiver of messages, an agent needs to choose whether or not to believe the received message(s). To help agents make this decision, this thesis introduces a Trust Model. In the trust model, receiver's impression on the sender, sender's reputation, and receiver's attitude towards risk are used to derive the receiver's trustworthiness on the sender. This thesis proposes that in making decisions on whether to believe a message and change the action based on the message, an agent should compare the persuasiveness of the message, which is calculated from the risk attitude of the receiver, the receiver's trustworthiness on the sender, and the utility brought by believing the message, with the stubbornness of the receiver. On the other hand, being a sender of messages, an agent needs to choose whether or not to be honest. To help agents make this decision, this thesis introduces a Honesty Model. In the honesty model, a sender uses its impression on the receiver, the receiver's reputation, and sender's

attitude towards risk to calculate the deceivability of a receiver. To decide whether to tell a lie, a sender compares the temptation of lying, which is derived from the sender's risk attitude, the receiver's deceivability, and the utility that will be gained by lying, with the sincerity of the sender. These mimic the model in human interactions.

In addition, we introduce an adaptive strategy to the Trust/Honesty Model, which enables agents to learn from and adapt to the environment. With the adaptive strategy, a receiver learns to be less risk-seeking and more stubborn after it is cheated. On the other hand, it learns to be more risk-seeking and less stubborn if it has not believed any messages for a long time or it has believed the right messages for many times. Similarly, a sender learns to be less risk-seeking and more sincere if it cannot gain the target receiver's trust. In contrast, it learns to be more risk-seeking and less sincere if it has not sent out any messages for a long time or it can successfully gain the receivers' trust for many times.

Simulations show that agents with the Adaptive Trust/Honesty Model perform much better than agents with other existing models or strategies. This is because our Adaptive Trust/Honesty Model enables agents to learn from their experiences. Another reason for the outstanding performance is that the Trust/Honesty Model makes a balance on trustworthiness and utility, while other existing models or strategies consider either trustworthiness or utility, but not both, in making the decisions.

在多代理人的環境裡，代理人需要選擇它們的行動。要達到這一點，Gmytrasiewicz 和 Durfee 提出遞歸塑造法。這份論文透過遞歸公式改進了原來設計的機率略計方式，這使代理人能預計其它代理人的行動和更加準確地選擇它們自己的行動。

在多代理人半競爭環境裡，代理人需要選擇其它類型的行動。在這樣的環境裡，競爭和合作是能夠並存的。當代理人互相競爭時，它們有誘因說謊。有些時候，代理人能夠透過互相合作來增加它們的利益，它們便會講真相。所以，作爲訊息接收者，代理人需要選擇是否相信接收到的訊息。爲幫助代理人做出這決策，這份論文介紹了信任模型。信任模型利用接收者對發送者的印象、發送者的名聲以及接收者對風險的態度來計算發送者的可信性。這份論文提出代理人在做出是否相信收到的訊息和根據收到的訊息更改行動這些決策時，應該計算自己對風險的態度、發送者的可信性以及訊息帶來的利益，作爲訊息的說服力，並且根據自己的倔強程度作出決策。另一方面，作爲訊息發送者，代理人需要選擇是否誠實。爲幫助代理人做出這決策，這份論文還介紹了誠實模型。誠實模型利用發送者對接收者的印象、接收者的名聲以及發送者對風險的態度來計算接收者是否容易受騙。在決定是否說謊，發送者計算自己對風險的態度、接收者的容易受騙程度以及說謊帶來的利益，作爲說謊的誘惑，並且根據自己的真誠度作出決策。這些仿造了人類的交往模式。

在信任／誠實模型上，我們還介紹了一個適應的方法，這使代理人能夠學習和適應環境。透過這個適應的方法，接收者學會在被欺詐之後變得較不愛風險和

更加倔強。另一方面，接收者學會在長期不相信任何消息或許多次相信了正確的消息之後變得較愛風險和較不倔強。同樣，發送者學會在無法獲取目標接收者的信任下變得較不愛風險和更加懇切。相反，發送者學會在長期未派出任何消息或許多次成功獲取接收者的信任後變得更愛風險和較不懇切。

試驗顯示，採用可適應的信任／誠實模型的代理人比採用其它現有的設計或方法的代理人有更好的表現。這是因爲我們這個可適應的信任／誠實模型使代理人能夠從他們的經驗學習。造成這優秀表現的其它原因是信任／誠實模型在可信性和利益之間做了一個平衡的考慮，但其它現有的設計或方法在決策中只考慮可信性或利益其中一頂。

Acknowledgments

This thesis would not be possible to completed without help of many people. I would like to take this opportunity to thank them.

Firstly, I would like to thank my Master Degree supervisor, Prof. Leung Ho Fung, for his guidance and encouragement in the past two years. He show his generosity and gave me numerous ideas for my research work.

I would like to thank Prof. Ng Wai Yin for suggestions and comments for improving this work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Autonomous agents and multiagent systems is one of the fields of specialization in artificial intelligence. According to Wooldridge and Jennings [WJ95]: *"An agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives"*. By adding flexibility to agents, agents become *intelligent agents*, also known as *autonomous agents*. Flexibility means reactivity, pro-activeness, and social ability. For agents to be flexible or intelligent, agents need to be able to react to the changes in the environment instantly, so as to achieve the design objectives. At the same time, intelligent agents need to be able to activate themselves and design their plans or choose their actions, in order to meet their goals. Adding to these, intelligent agents need to have social ability, which is the ability to cooperate or negotiate with other agents.

Another concept closely related to intelligence is learning. An agent is said to have added intelligence if it is able to learn, which means the agent is able to improve its future behaviors based on its past experiences. There are two principal categories of learning in multiagent systems [SW00]. First is *centralized learning*, also known as isolated learning. With this kind of learning, agents learn by itself, and independent of other agents. The second category is *decentralized learning*, also known as interactive learning. With interactive learning, agents learn from each other. Agents learn by exchanging information with each other, that is through communication. Agents can learn by simple query-and-answer interactions, or more complex interactions, such as negotiation.

Communications not only enable agents to learn, but also enable agents to coordinate their actions. In multiagent systems, agents may have different objectives, actions and behaviors. In order to achieve the design objectives, agents need to coordinate their actions. There are two types of coordination, which are cooperation and competition. In cooperative environments, agents need to have distributed or centralized planning for their actions. In competitive or non-cooperative environments, agents need to negotiate to resolve conflicts. In multiagent systems, agent coordination has been the subject of continuous interest. Much research has been done on protocols and modeling. In different environments, agents have different methods or protocols to resolve conflicts. In non-cooperative environments, Zlotkin and Rosenschein [ZR90] introduce a theoretical negotiation model, which encompasses both cooperative and conflicting situations. Besides, they also use a conflict resolution protocol to help agents reach agreement. One the other hand, they use another negotiation protocol [ZR89] to help agents share their tasks in cooperative environments, so that agents can communicate their respective desires and compromise to reach mutually beneficial agreements. At the same time, Rosenschein and Genesereth [RG85] use a deal-making mechanism to enable agents to cooperate. Through the use of communication and binding promises, agents are able to coordinate their actions effectively. This also makes mutually beneficial activities possible. In contrast to the pre-established protocols mentioned above, Gmytrasiewicz and Durfee [GD95, GD00, GD01] propose a decision-theoretic approach [Bra92, Rai82, GD00], called the Recursive Modeling Method (RMM), which enables agents to choose an action rationally in the absence of any conventions.

Trust and reputation is a hot topic in agent coordination. In the literature, there are different meanings for "trusting an agent". Some interpret "trusting an agent" to be "cooperate with an agent" [Mar94, MMH02, MMA+01, MHM02], while others interpret that as "delegate to an agent" [CF98, FC01]. At the same time, there are various models and definitions for trust and reputation [Mar94, MMH02, MMA+01, MHM02, SS01, RLM01, YS01, YS03, CF98, FC01, GG00]. Marsh [Mar94] relates trust to the risk and importance of a matter, as well as the competence of a particular agent on that particular matter. Mui *et al.* [MMH02, MMA+01] define trust as the expected probability that an agent will cooperate the next time, given a history of encounters. For reputation, various definitions [MHM02, SS01, RLM01, YS01, YS03] are similar to a weighted sum of individual experience.

1.1 Motivations

For agents to coordinate in cooperative environments, Gmytrasiewicz and Durfee [GD95, GD00, GD01] propose the Recursive Modeling Method (RMM). RMM is a decision-theoretic approach [Bra92, Rai82, GD00], which enables agents to choose an action rationally in the absence of any conventions. In general, rationality means the maximization of expected utility [Fis81]. This approach uses RMM to represent the information that an agent has about the environment, itself, as well as other agents. This information enables agents to predict the actions of other agents, which in turn helps agents choose their own actions. However, the authors use a probabilistic approximation in their design, which may introduce inaccuracy.

For trust and honesty, much research has been done on cooperative environments. In purely cooperative environments, benevolent agents share their utilities as social welfare. As agents have a common goal of maximizing the social welfare, there is no reason for an agent to be dishonest to its partners. So, there is no reason for an agent not to trust its partners. On the other hand, in strictly competitive environments (such as zero-sum games), only one agent can be the winner and the others must be the losers. As agents are self-interested and cannot increase their utilities by cooperating with each other, it is rational for an agent to be dishonest. So, it is irrational for an agent to believe information provided by its competitors. Luo *et al.*

[LJS+03] define semi-competitive environment to be an environment having both cooperations and competitions, in which agents seek to strike a fair deal for both parties, and at the same time try to maximize their own payoffs. Therefore, it is sometimes rational for an agent to cooperate with some other agents, while agents also have incentives to be dishonest. Therefore, being a receiver, an agent needs to decide whether or not to trust other agents on receiving information from other agents. On the other hand, being a sender, an agent needs to decide whether or not to lie to other agents. In such a setting, the issues about trust and honesty among agents become more significant and complicated than those in purely cooperative or strictly competitive environments.

1.2 Aims

There are two aims in this thesis. The first one is to improve the original design of Gmytrasiewicz and Durfee's Recursive Modeling Method. To improve their probabilistic approximation, we introduce our Recursive Formulas. This is to help agents predict other agents' actions and choose their own actions.

Another aim is to develop a Trust/Honesty Model, which helps agents choose their communication actions. This model helps receivers choose whether to believe a received message, and choose which message to believe when several messages are received. This model also helps senders choose whether to tell lies.

1.3 Contributions

First, we improve the probabilistic approximation in the original RMM by Recursive Formulas. This enables agents to predict other agents' actions and choose their own actions more accurately.

Second, we develop a Trust/Honesty Model, which helps agents choose their communication actions. This enables receivers to choose whether or not to trust the received information, and enables senders to choose whether or not to be honest in semi-competitive environments.

In addition, we develop an adaptive strategy for the Trust/Honesty Model. This enables agents to learn from their experiences and to adapt to the interacting agents. The adaptive agent we built is an intelligent agent, which can react to its opponents, maximize its payoff actively, and able to interact with other agents. Simulations show that the adaptive agent performs much better than agents with existing models or strategies.

1.4 Thesis Outline

The rest of the thesis is organized as follows.

In the next chapter, we describe the improved Recursive Modeling Method. First, we introduce the original RMM with an illustrative example. Then, we improve the original design by Recursive Formulas. Last, we compare the original RMM with the modified one.

In Chapter 3, we present our Trust/Honesty Model. We first show the needs for such a model. Then we present our model.

In Chapter 4, we present an adaptive strategy for the Trust/Honesty Model. First, we point out the problem with the non-adaptive agents. Then, we introduce the adaptive strategy. Last, we compare the adaptive agents with the non-adaptive ones.

In Chapter 5, we discuss the related work and point out the problems with the existing models.

In Chapter 6, we analyze the performance of our Trust/Honesty Model and the adaptive strategy. Simulations are done to compare performance of agents adopting our Trust/Honesty Model with/without the adaptive strategy with agents adopting other existing models and strategies. Performance of agents in semi-competitive environment and performance of agents when interacting with strategic senders are analyzed.

Chapter 7 concludes the paper and discusses some possible future work.

Chapter 2

Improved Recursive Modeling Method (RMM)

RMM is a method for agents to represent their knowledge and choose their actions. This chapter introduces the background knowledge about Recursive Modeling Method (RMM) with an illustrative example. In section 2.2, we give a brief introduction to the original RMM. After that, we give a detailed description on an improved RMM in section 2.3. Finally, we compare the original RMM with the improved version.

2.1 An Illustrative Example

Consider an example of agent interaction. In this example, there are three agents: R_1, R_2 and R_3, and there are three goals: G_1, G_2 and G_3. The example scenario is depicted in Fig. 2.1.

For all agents R_i, where i = 1, 2, 3, in the environment, they share the same set of possible actions: obtaining G_1, obtaining G_2, obtaining G_3, or staying still, which are denoted as A = $\{R_i \rightarrow G_1, R_i \rightarrow G_2, R_i \rightarrow G_3, R_i \rightarrow S\}$. However, the set of actions A_i for agent i may only be a subset of A, depending on the set of goals that the agent knows. Knowledge of the agents is shown in Table 2.1. For example, R_2 does not know G_3, so A_2 will be $\{R \rightarrow G_1, R \rightarrow G_2, R \rightarrow S\}$ because obtaining

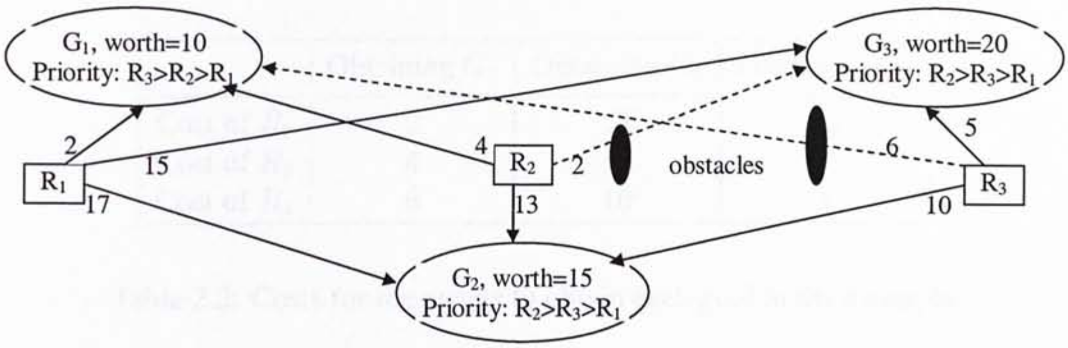

Figure 2.1: Example scenario of interacting agents

R_1's knowledge	R_2's knowledge	R_3's knowledge
G_1	G_1	G_2
G_2	G_2	G_3
G_3		
R_2 does not know G_3		
R_3 does not know G_1		

Table 2.1: Agents' knowledge in the example

G_3 will not be an option for R_2 as it does not know that goal. We assume that if an agent knows a certain goal, the agent knows all the relevant information of the goal.[1]

In each round, agents communicate and choose their actions. After an agent has chosen its action, it needs to pay the cost so as to obtain the goal. The costs for the agents to obtain the goals are shown in Table 2.2, and the cost for staying still ($R\rightarrow S$) is zero. The payoff for R_i to obtain G_j is defined to be the worth of the goal (if R_i wins the worth of G_j) minus the cost for the agent to obtain the goal. In this way, if R_i cannot win the worth of the goal, its payoff will be negative. For example, if R_1 chooses to obtain G_1 and it wins, its payoff will be $10 - 2 = 8$. However, if it losses, its payoff will be -2.

It is possible that more than one agent may choose to obtain the same goal. In

[1]It is possible that an agent knows only some information of a goal while not knowing others. We do not consider this as this will much complicate the discussion.

	Obtaining G_1	Obtaining G_2	Obtaining G_3
Cost of R_1	2	17	15
Cost of R_2	4	13	2
Cost of R_3	6	10	5

Table 2.2: Costs for the agents to obtain each goal in the example

	G_1	G_2	G_3
1^{st} priority	R_3	R_2	R_2
2^{nd} priority	R_2	R_3	R_3
3^{rd} priority	R_1	R_1	R_1

Table 2.3: Goal's priority ordering in the example

this case, the worth of the goal will be completely given to one agent among all the full-cost paying agents. To decide which agent can win the worth of a certain goal, different systems may apply different mechanisms. For ease of presentation and without loss of generality, we assume that for each goal, there is an associated priority ordering of agents such that when more than one agent decides to obtain the same goal, the worth of the goal will be given to the agent according to the goal's priority ordering of agents. The priority ordering of the goals in the example are summarized in Table 2.3. For example, if all the three agents decide to obtain G_1, R_3 will win the worth of G_1. If R_3 decides not to obtain G_1, and both R_2 as well as R_1 decide to obtain G_1, R_2 will win the worth. If R_1 decides to obtain G_1, it can win the worth only if both R_3 and R_2 do not compete with it.

We define a *Single-round Game* to be a game consists of only one round. In a Single-round Game, agents are free to communicate until all agents openly announce their choices of actions and pay the costs, then the game ends and the worths of the goals are given to the winning agents. Each agent can only take one action. In this chapter, we illustrate how an agent chooses its action in a Single-round Game.

2.2 Recursive Modeling Method (RMM)

For agents to represent the above information and take the appropriate action, Gmytrasiewicz and Durfee [GDW91, GD92, GD95, GD00, GD01] propose the Recursive Modeling Method (RMM). RMM represents the information an agent has about the environment, itself, and other agents. This representation is used by an agent to predict the actions of other agents, estimate the expected utilities for alternative courses of action, and make decisions on what actions or communication acts to perform. RMM is recursive as it not only represents an agent's own preferences, abilities and beliefs about the world, but also represents the beliefs the agent has about other agents, the beliefs it has about other agents' beliefs, and so on. So it is basically an infinite hierarchy. However, Gmytrasiewicz and Durfee have made an assumption that the belief hierarchy is finite and terminates at the point where an agent has no sufficient information to model other agents. At the point of insufficient information, the infinite belief hierarchy is terminated to a finite one by an assumed uniform distribution over the space of all possible actions.

2.2.1 Payoff Matrices

With RMM, agents' payoffs are represented in payoff matrices. In this three-agent environment, we use a three-dimensional payoff matrix in modeling agent's decision-making process. Fig. 2.2 shows an example of a cell. The cell is a two-dimensional payoff matrix describing R_1's payoffs with respect to its own actions, R_2's actions and R_3's action to obtain G_1. For example, $p^{R_1 \to G_2}_{R_2 \to G_3, R_3 \to G_1}$ denotes the payoff for R_1 to obtain G_2 when R_2 chooses to obtain G_3 and R_3 chooses the obtain G_1.

From Table 2.1, R_1 knows all the goals, so it has four possible actions. With the information in Table 2.3 and Table 2.2, if both R_2 and R_3 choose to obtain G_1 and R_1 also chooses to obtain G_1, since R_1 has the lowest priority to get the worth of G_1, it will lose and its payoff will be zero minus its cost to obtain G_1, which is

$R_3\to G_1$				
	$R_2\to G_1$	$R_2\to G_2$	$R_2\to G_3$	$R_2\to S$
$R_1\to G_1$	$p^{R_1\to G_1}_{R_2\to G_1,R_3\to G_1}$	$p^{R_1\to G_1}_{R_2\to G_2,R_3\to G_1}$	$p^{R_1\to G_1}_{R_2\to G_3,R_3\to G_1}$	$p^{R_1\to G_1}_{R_2\to S,R_3\to G_1}$
$R_1\to G_2$	$p^{R_1\to G_2}_{R_2\to G_1,R_3\to G_1}$	$p^{R_1\to G_2}_{R_2\to G_2,R_3\to G_1}$	$p^{R_1\to G_2}_{R_2\to G_3,R_3\to G_1}$	$p^{R_1\to G_2}_{R_2\to S,R_3\to G_1}$
$R_1\to G3$	$p^{R_1\to G_3}_{R_2\to G_1,R_3\to G_1}$	$p^{R_1\to G_3}_{R_2\to G_2,R_3\to G_1}$	$p^{R_1\to G_3}_{R_2\to G_3,R_3\to G_1}$	$p^{R_1\to G_3}_{R_2\to S,R_3\to G_1}$
$R_1\to S$	$p^{R_1\to S}_{R_2\to G_1,R_3\to G_1}$	$p^{R_1\to S}_{R_2\to G_2,R_3\to G_1}$	$p^{R_1\to S}_{R_2\to G_3,R_3\to G_1}$	$p^{R_1\to S}_{R_2\to S,R_3\to G_1}$

Figure 2.2: An example of a cell in a three-dimensional matrix

$P^{R_1\text{-}R_1}$

$R_3\to G_1$				
	$R_2\to G_1$	$R_2\to G_2$	$R_2\to G_3$	$R_2\to S$
$R_1\to G_1$	-2	-2	-2	-2
$R_1\to G_2$	-2	-17	-2	-2
$R_1\to G_3$	5	5	-15	5
$R_1\to S$	0	0	0	0

$R_3\to G_2$				
	$R_2\to G_1$	$R_2\to G_2$	$R_2\to G_3$	$R_2\to S$
$R_1\to G_1$	-2	8	8	8
$R_1\to G_2$	-17	-17	-17	-17
$R_1\to G_3$	5	5	-15	5
$R_1\to S$	0	0	0	0

$R_3\to G_3$				
	$R_2\to G_1$	$R_2\to G_2$	$R_2\to G_3$	$R_2\to S$
$R_1\to G_1$	-2	8	8	8
$R_1\to G_2$	-2	-17	-2	-2
$R_1\to G_3$	-15	-15	-15	-15
$R_1\to S$	0	0	0	0

$R_3\to S$				
	$R_2\to G_1$	$R_2\to G_2$	$R_2\to G_3$	$R_2\to S$
$R_1\to G_1$	-2	8	8	8
$R_1\to G_2$	-2	-17	-2	-2
$R_1\to G_3$	5	5	-15	5
$R_1\to S$	0	0	0	0

Figure 2.3: R_1's payoff matrix

$0 - 2 = -2$. In another situation, if both R_2 and R_3 choose to obtain G_2 and R_1 chooses to obtain G_1, since R_1 has no competitor in this case, its payoff will be the worth of G_1 minus its cost to obtain G_1, which is $10 - 8 = 2$. In this way, R_1 represents its knowledge in its own payoff matrix, which is shown in Fig. 2.3.

In addition, R_1 can model R_2's payoff matrix and R_3's payoff matrix from their respective points of view, which are shown in Fig. 2.4 and Fig. 2.5. $P^{R_1-R_2}$ denotes R_1's model of R_2's payoff matrix, while $P^{R_1-R_3}$ denotes R_1's model of R_3's payoff matrix. From Table 2.1, R_1 knows that R_2 does not know G_3 and R_3 does not know G_1. As R_1 knows that R_2 does not know G_3, there will not be any payoffs associate to G_3 in R_2's payoff matrix. Similarly, there will not be any payoffs associate to G_1

$P^{R_1-R_2}$

$R_3 \rightarrow G_1$				$R_3 \rightarrow G_2$			
	$R_1 \rightarrow G_1$	$R_1 \rightarrow G_2$	$R_1 \rightarrow S$		$R_1 \rightarrow G_1$	$R_1 \rightarrow G_2$	$R_1 \rightarrow S$
$R_2 \rightarrow G_1$	-4	-4	-4	$R_2 \rightarrow G_1$	6	6	6
$R_2 \rightarrow G_2$	2	2	2	$R_2 \rightarrow G_2$	2	2	2
$R_2 \rightarrow S$	0	0	0	$R_2 \rightarrow S$	0	0	0
				$R_3 \rightarrow S$			
					$R_1 \rightarrow G_1$	$R_1 \rightarrow G_2$	$R_1 \rightarrow S$
				$R_2 \rightarrow G_1$	6	6	6
				$R_2 \rightarrow G_2$	2	2	2
				$R_2 \rightarrow S$	0	0	0

Figure 2.4: R_1's model of R_2's payoff matrix

$P^{R_1-R_3}$

				$R_1 \rightarrow G_2$			
					$R_2 \rightarrow G_2$	$R_2 \rightarrow G_3$	$R_2 \rightarrow S$
				$R_3 \rightarrow G_2$	-10	5	5
				$R_3 \rightarrow G_3$	-5	15	15
				$R_3 \rightarrow S$	0	0	0
$R_1 \rightarrow G_3$				$R_1 \rightarrow S$			
	$R_2 \rightarrow G_2$	$R_2 \rightarrow G_3$	$R_2 \rightarrow S$		$R_2 \rightarrow G_2$	$R_2 \rightarrow G_3$	$R_2 \rightarrow S$
$R_3 \rightarrow G_2$	-10	5	5	$R_3 \rightarrow G_2$	-10	5	5
$R_3 \rightarrow G_3$	-5	15	15	$R_3 \rightarrow G_3$	-5	15	15
$R_3 \rightarrow S$	0	0	0	$R_3 \rightarrow S$	0	0	0

Figure 2.5: R_1's model of R_3's payoff matrix

in R_3's payoff matrix.

2.2.2 Infinite Recursive Hierarchy

With RMM, for R_1 to determine its action, it first models how R_2 and R_3 determine their actions. For R_1 to deduce R_2's action, it has to model how R_2 deduce the actions of R_1 and R_3. Similarly, for R_1 to deduce R_3's action, it has to model how R_3 deduce the actions of R_1 and R_2, and so on. This generates an infinite recursive hierarchy of payoff matrices as shown in Fig. 2.6. At the first level, there is R_1's model of itself, denoted as $P^{R_1-R_1}$. At the second level, there are R_1's models of

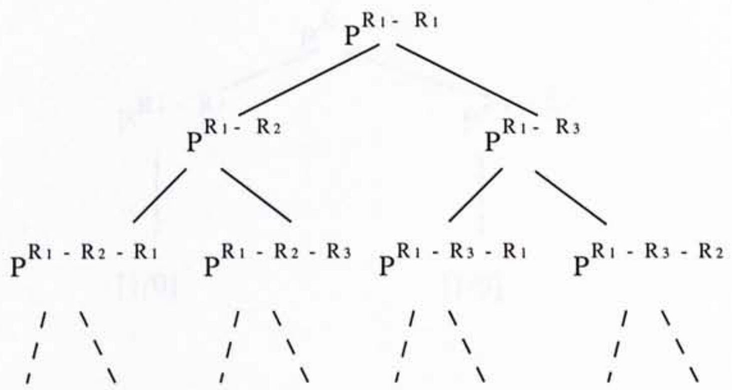

Figure 2.6: R_1's infinite recursive hierarchy

R_2's and R_3's knowledge, denoted as $P^{R_1-R_2}$ and $P^{R_1-R_3}$, respectively. At the third level, there are, from R_1's point of view, R_2's models of R_1's and R_3's knowledge, as well as R_3's models of R_1's and R_2's knowledge. Similarly, the hierarchy goes on infinitely.

2.2.3 Choosing an Action with RMM

For R_1 to deduce the payoff matrices at the third level, that is to model R_2's models of R_1's and R_3's knowledge, as well as R_3's models of R_1's and R_2's knowledge, R_1 needs further information. For example, R_1 has to know whether R_2 knows that R_1 knows G_1 or whether R_2 knows that R_3 does not know G_1. Since R_1 has no further information on R_2's knowledge of other agents' knowledge, in the original design of RMM, R_1's infinite recursive hierarchy is terminated explicitly to a finite one at this level, which is shown in Fig. 2.7. At the first level, it is R_1's model of itself. At the second level, there are R_1's models of R_2's and R_3's payoff matrices. Since R_1 knows that R_2 does not know G_3, there are only three possible actions from R_2's point of view, which results in nine combinations of actions for the three agents. So, at the third level, the hierarchy on the left is terminated with a uniform probability distribution: $\frac{1}{9}$ for each possible action. The hierarchy on the right terminates in

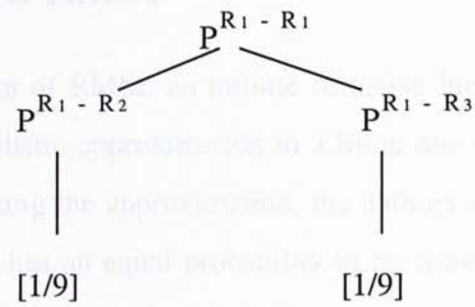

Figure 2.7: R_1's finite recursive hierarchy with original design of RMM

Action	Expected utilities
$R_2 \rightarrow G_1$	$\frac{1}{9} \times p^{R_2 \rightarrow G_1}_{R_1 \rightarrow G_1, R_3 \rightarrow G_1} + \frac{1}{9} \times p^{R_2 \rightarrow G_1}_{R_1 \rightarrow G_2, R_3 \rightarrow G_1} + \frac{1}{9} \times p^{R_2 \rightarrow G_1}_{R_1 \rightarrow S, R_3 \rightarrow G_1} + \frac{1}{9} \times p^{R_2 \rightarrow G_1}_{R_1 \rightarrow G_1, R_3 \rightarrow G_2} + \frac{1}{9} \times p^{R_2 \rightarrow G_1}_{R_1 \rightarrow G_2, R_3 \rightarrow G_2} + \frac{1}{9} \times p^{R_2 \rightarrow G_1}_{R_1 \rightarrow S, R_3 \rightarrow G_2} + \frac{1}{9} \times p^{R_2 \rightarrow G_1}_{R_1 \rightarrow G_1, R_3 \rightarrow S} + \frac{1}{9} \times p^{R_2 \rightarrow G_1}_{R_1 \rightarrow G_2, R_3 \rightarrow S} + \frac{1}{9} \times p^{R_2 \rightarrow G_1}_{R_1 \rightarrow S, R_3 \rightarrow S}$ $= \frac{1}{9} \times -4 + \frac{1}{9} \times -4 + \frac{1}{9} \times -4 + \frac{1}{9} \times 6 + \frac{1}{9} \times 6 + \frac{1}{9} \times 6 + \frac{1}{9} \times 6 + \frac{1}{9} \times 6 + \frac{1}{9} \times 6 = \frac{24}{9}$
$R_2 \rightarrow G_2$	$\frac{1}{9} \times p^{R_2 \rightarrow G_2}_{R_1 \rightarrow G_1, R_3 \rightarrow G_1} + \frac{1}{9} \times p^{R_2 \rightarrow G_2}_{R_1 \rightarrow G_2, R_3 \rightarrow G_1} + \frac{1}{9} \times p^{R_2 \rightarrow G_2}_{R_1 \rightarrow S, R_3 \rightarrow G_1} + \frac{1}{9} \times p^{R_2 \rightarrow G_2}_{R_1 \rightarrow G_1, R_3 \rightarrow G_2} + \frac{1}{9} \times p^{R_2 \rightarrow G_2}_{R_1 \rightarrow G_2, R_3 \rightarrow G_2} + \frac{1}{9} \times p^{R_2 \rightarrow G_2}_{R_1 \rightarrow S, R_3 \rightarrow G_2} + \frac{1}{9} \times p^{R_2 \rightarrow G_2}_{R_1 \rightarrow G_1, R_3 \rightarrow S} + \frac{1}{9} \times p^{R_2 \rightarrow G_2}_{R_1 \rightarrow G_2, R_3 \rightarrow S} + \frac{1}{9} \times p^{R_2 \rightarrow G_2}_{R_1 \rightarrow S, R_3 \rightarrow S}$ $= \frac{1}{9} \times 2 + \frac{1}{9} \times 2 + \frac{1}{9} \times 2 + \frac{1}{9} \times 2 + \frac{1}{9} \times 2 + \frac{1}{9} \times 2 + \frac{1}{9} \times 2 + \frac{1}{9} \times 2 + \frac{1}{9} \times 2 = 2$
$R_2 \rightarrow S$	0

Table 2.4: R_2's expected utilities

a similar way. Here, the authors of the original RMM assume that each possible action has an equal probability to be chosen.

With this hierarchy, from R_1's model of R_2's payoff matrix shown in Fig. 2.4, R_1 calculates R_2's expected utilities for each action, which is shown in Table 2.4. From this, since the expected utilities for R_2 to obtain G_1 is the highest, R_1 models that R_2 will choose to obtain G_1. Similarly, R_1 models that R_3 will choose to obtain G_3. If R_2 chooses to obtain G_1 and R_3 chooses to obtain G_3, from R_1's payoff matrix in Fig. 2.3, R_1 can only choose to stay still.

2.3 Improved RMM

In the original design of RMM, an infinite recursive hierarchy is terminated explicitly by a probabilistic approximation to a finite one when there is no further information. In making the approximation, the authors made an assumption that each possible action has an equal probability to be chosen. However, this is not always accurate in real practice. In this section, we present our improved version of RMM, which solves the problem.

2.3.1 Infinite Recursive Hierarchy

Suppose R_1 only has level-1 knowledge, which means that besides the set of goals that it knows, the agent does not know other agents' knowledge. In this case, it can only assume other agents know the same set of goals as it does. As the whole hierarchy is built from R_1's point of view, the set of actions involved in the payoff matrices in the hierarchy will be the set of actions as seen by R_1, which is A_1. R_1's infinite recursive hierarchy with level-1 knowledge is shown in Fig. 2.8. The hierarchy is basically the same as the one shown in Fig. 2.6, but as R_1 only has level-1 knowledge, all the payoff matrices in this hierarchy are constructed with R_1's set of actions. Note that $P_{A_1}^{R_1-R_1}$ is the same as $P^{R_1-R_2}$. However, if R_1 does not know that R_2 does not know G_3 and R_3 does not know G_1, it is rational for it to model their payoff matrices with its own set of possible actions, which is A_1. In this case, R_1's models of R_2's and R_3's payoff matrices are shown in Fig. 2.9 and Fig. 2.10.

2.3.2 The Sub-matrix Operator

In the example, R_1 has level-2 knowledge in addition to level-1 knowledge, which means that in addition to the set of goals that it knows, it also knows the set of goals that R_2 and R_3 know, as well as the set of goals that they don't know. Note that

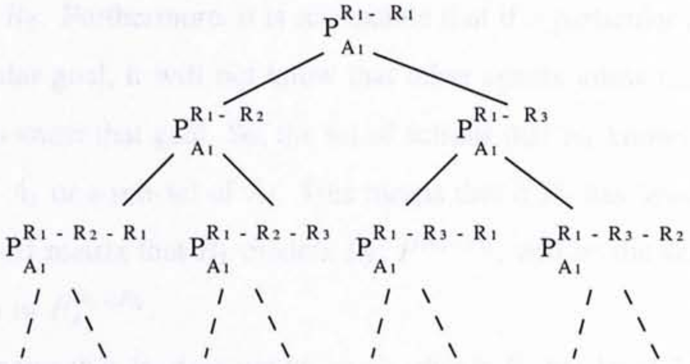

Figure 2.8: R_1's infinite recursive hierarchy with level-1 knowledge

$P_{A_1}^{R_1-R_2}$

	$R_3 \to G_1$					$R_3 \to G_2$			
	$R_1 \to G_1$	$R_1 \to G_2$	$R_1 \to G_3$	$R_1 \to S$		$R_1 \to G_1$	$R_1 \to G_2$	$R_1 \to G_3$	$R_1 \to S$
$R_2 \to G_1$	-4	-4	-4	-4	$R_2 \to G_1$	6	6	6	6
$R_2 \to G_2$	2	2	2	2	$R_2 \to G_2$	2	2	2	2
$R_2 \to G_3$	18	18	18	18	$R_2 \to G_3$	18	18	18	18
$R_2 \to S$	0	0	0	0	$R_2 \to S$	0	0	0	0
	$R_3 \to G_3$					$R_3 \to S$			
	$R_1 \to G_1$	$R_1 \to G_2$	$R_1 \to G_3$	$R_1 \to S$		$R_1 \to G_1$	$R_1 \to G_2$	$R_1 \to G_3$	$R_1 \to S$
$R_2 \to G_1$	6	6	6	6	$R_2 \to G_1$	6	6	6	6
$R_2 \to G_2$	2	2	2	2	$R_2 \to G_2$	2	2	2	2
$R_2 \to G_3$	18	18	18	18	$R_2 \to G_3$	18	18	18	18
$R_2 \to S$	0	0	0	0	$R_2 \to S$	0	0	0	0

Figure 2.9: R_1's model of R_2's payoff matrix with set of actions A_1

$P_{A_1}^{R_1-R_3}$

	$R_1 \to G_1$					$R_1 \to G_2$			
	$R_2 \to G_1$	$R_2 \to G_2$	$R_2 \to G_3$	$R_2 \to S$		$R_2 \to G_1$	$R_2 \to G_2$	$R_2 \to G_3$	$R_2 \to S$
$R_3 \to G_1$	4	4	4	4	$R_3 \to G_1$	4	4	4	4
$R_3 \to G_2$	5	-10	5	5	$R_3 \to G_2$	5	-10	5	5
$R_3 \to G_3$	15	-5	15	15	$R_3 \to G_3$	15	-5	15	15
$R_3 \to S$	0	0	0	0	$R_3 \to S$	0	0	0	0
	$R_1 \to G_3$					$R_1 \to S$			
	$R_2 \to G_1$	$R_2 \to G_2$	$R_2 \to G_3$	$R_2 \to S$		$R_2 \to G_1$	$R_2 \to G_2$	$R_2 \to G_3$	$R_2 \to S$
$R_3 \to G_1$	4	4	4	4	$R_3 \to G_1$	4	4	4	4
$R_3 \to G_2$	5	-10	5	5	$R_3 \to G_2$	5	-10	5	5
$R_3 \to G_3$	15	-5	15	15	$R_3 \to G_3$	15	-5	15	15
$R_3 \to S$	0	0	0	0	$R_3 \to S$	0	0	0	0

Figure 2.10: R_1's model of R_3's payoff matrix with set of actions A_1

it is possible for R_1 to have level-2 knowledge on R_2 and has a different level of knowledge on R_3. Furthermore, it is reasonable that if a particular agent does not know a particular goal, it will not know that other agents know that goal even if other agents do know that goal. So, the set of actions that R_1 knows about R_2 will be the same as A_1 or a sub-set of A_1. This means that if R_1 has level-2 knowledge on R_2, the payoff matrix that R_1 models R_2, $P^{R_1-R_2}$, will be the same as $P_{A_1}^{R_1-R_2}$ or a sub-matrix of $P_{A_1}^{R_1-R_2}$.

Since R_1 knows that R_2 does not know G_3, that is R_1 has level-2 knowledge on R_2, its model of R_2's payoff matrix $P^{R_1-R_2}$, shown in Fig. 2.4, actually is equal to a sub-matrix of $P_{A_1}^{R_1-R_2}$, shown in Fig. 2.9, with rows and columns associated with G_3 removed.

Now, let us define a sub-matrix operator: $\ominus_{\mathcal{G}}$, such that if $P' = P\ominus_{\mathcal{G}}$, then P' will be equal to P, with rows and columns associated with the set of goals $\mathcal{G}$ removed. So in the example, R_1's model of R_2's payoff matrix with level-2 knowledge will be denoted as $P_{A_1}^{R_1-R_2}\ominus_{\{G_3\}}$.

This sub-matrix operator will have the following property:

$$P\ominus_{\mathcal{G}_a}\ominus_{\mathcal{G}_b} = P\ominus_{\mathcal{G}_b}\ominus_{\mathcal{G}_a} = P\ominus_{\mathcal{G}_a\cup\mathcal{G}_b} = P\ominus_{\mathcal{G}_b\cup\mathcal{G}_a},$$

where $P\ominus_{\mathcal{G}_a\cup\mathcal{G}_b}$ means the resulting sub-matrix will be equal to P, with all rows and columns associated with both sets of goals $\mathcal{G}_a$, and $\mathcal{G}_b$ removed.

2.3.3 Finite Recursive Hierarchy

We introduce another notation $\mathcal{G}_{ij}$, which denotes the set of goals that R_i knows R_j does not know, and $\mathcal{G}_{ijk}$, which denotes the set of goals that R_i knows R_j knows R_k does not know, and so on. With the sub-matrix operator, R_1's infinite recursive hierarchy will become Fig. 2.11. At the second level, since R_1 knows R_2 does not know the set of goals $\mathcal{G}_{12}$, R_1 needs not include this set of goals in R_1's model of R_2's payoff matrix. So, $P^{R_1-R_2}$ will be a sub-matrix of $P_{A_1}^{R_1-R_2}$, with rows and columns associated with set of goals $\mathcal{G}_{12}$ removed. At the third level, since R_1

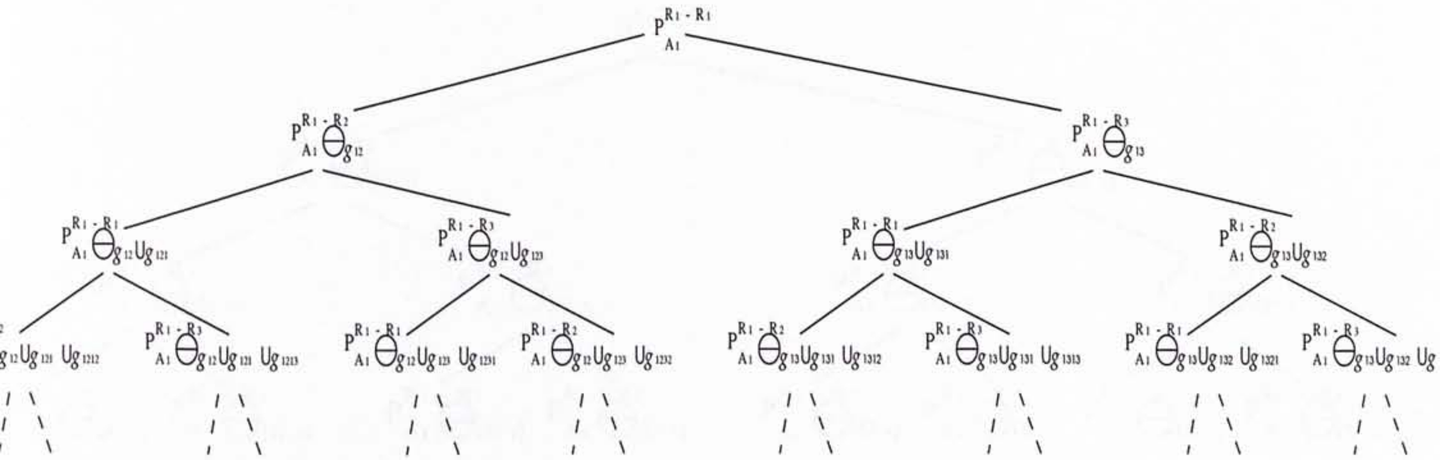

Figure 2.11: R_1's infinite recursive hierarchy with sub-matrix operator

knows R_2 does not know the set of goals $\mathcal{G}_{12}$, and R_1 knows R_2 knows R_1 does not know the set of goals $\mathcal{G}_{121}$, R_1 needs not include these sets of goals in its model of R_2's models of R_1's payoff matrix. So, $P^{R_1-R_2-R_1}$ will be a sub-matrix of $P_{A_1}^{R_1-R_1}$, with rows and columns associated with sets of goals in $\mathcal{G}_{12}$ and $\mathcal{G}_{121}$ removed, and so on.

In the example, R_1 has level-2 knowledge on both R_2 and R_3: "R_1 knows that R_2 does not know G_3" and "R_1 knows that R_3 does not know G_1". This means that $\mathcal{G}_{12} = \{G_3\}$, $\mathcal{G}_{13} = \{G_1\}$. Since R_1 does not have further knowledge, all $\mathcal{G}_{121}$, $\mathcal{G}_{123}$, $\mathcal{G}_{131}$, $\mathcal{G}_{132}$, $\mathcal{G}_{1212}$, $\mathcal{G}_{1213}$, $\mathcal{G}_{1231}$, $\mathcal{G}_{1232}$, $\mathcal{G}_{1312}$, $\mathcal{G}_{1313}$, $\mathcal{G}_{1321}$, $\mathcal{G}_{1323}$, and so on, will be empty sets. In this case, the recursive hierarchy can be further simplified to the one shown in Fig. 2.12. At the second level, since R_1 knows that R_2 does not know G_3, R_1 needs not include G_3 in its model of R_2's payoff matrix. At the third level, for R_1 to model R_2's model of R_1's payoff matrix, R_1 also needs not include G_3. At the forth level, for R_1 to model R_2's model of R_1's model of R_2's payoff matrix, again G_3 needs not be included. From the figure, we can see that since level-3 knowledge is not available for R_1, the payoff matrices at level 4 of the hierarchy cannot be reduced anymore, and starts to repeat the patterns at upper levels. For

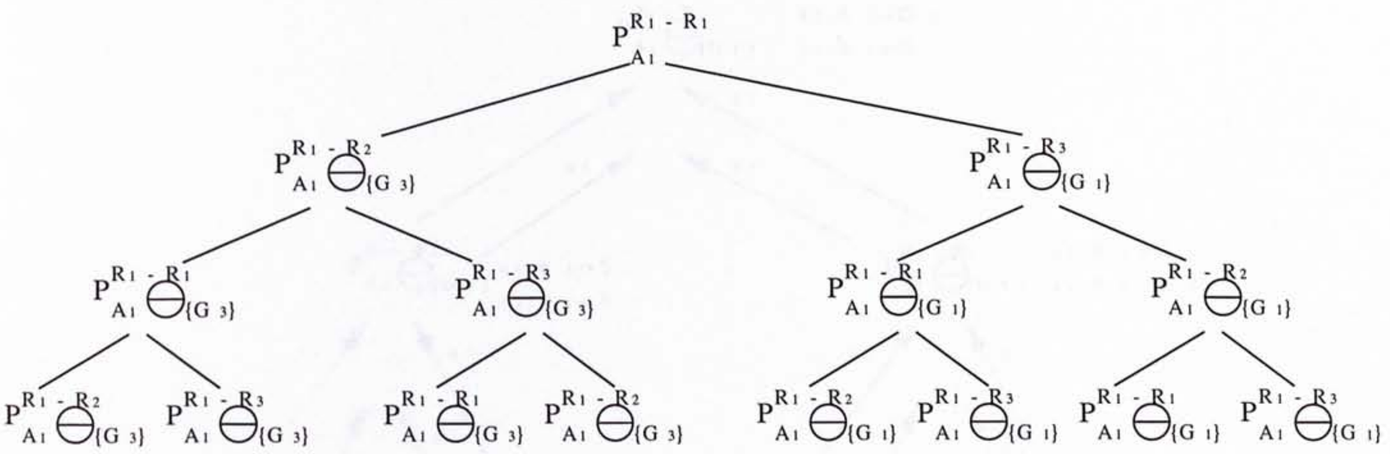

Figure 2.12: R_1's finite recursive hierarchy with level-2 knowledge

example, the payoff matrix $P_{A_1}^{R_1-R_2}\ominus_{\{G_3\}}$ at level four repeats that at level two, and the payoff matrix $P_{A_1}^{R_1-R_3}\ominus_{\{G_3\}}$ at level four repeats that at level three. So, the infinite recursive hierarchy can be terminated at this level, and becomes a finite hierarchy.

In general, an infinite recursive hierarchy can be terminated at level (*k+1*) if level-*k* knowledge is not available. Same as the original design of RMM, we terminate an infinite hierarchy due to lack of further information, which agents can use to model other agents.

2.3.4 Choosing an Action

From Fig. 2.12, we can see that the payoff matrices at the forth level repeat those at the second level and those at the third level. By method of iteration, an action can be chosen. Let's look at the left sub-tree of Fig. 2.12, as shown in Fig. 2.13. Starting from the payoff matrix on the left bottom corner in Fig. 2.13, which is $P_{A_1}^{R_1-R_2}\ominus_{\{G_3\}}$, we can deduce the best action for R_2 is a_1: $R_2 \rightarrow G_1$, which means R_2 can get the highest payoff if it chooses to obtain G_1. We apply a_1 to $P_{A_1}^{R_1-R_3}\ominus_{\{G_3\}}$, that is if R_2 chooses to obtain G_1, from $P_{A_1}^{R_1-R_3}\ominus_{\{G_3\}}$, we can deduce the best

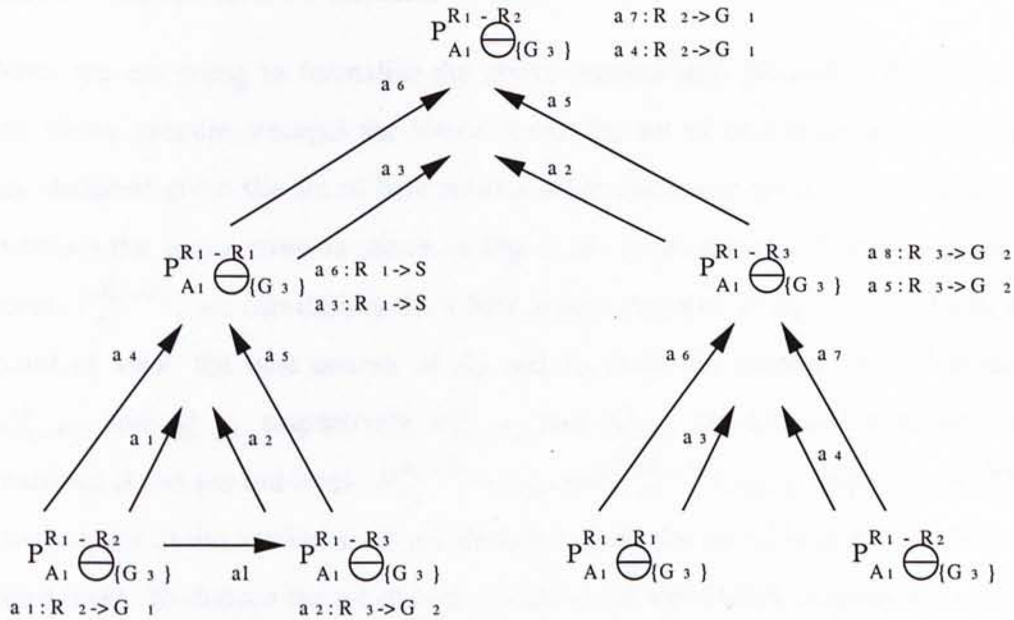

Figure 2.13: Left sub-tree of Fig. 2.12

action for R_3 is a_2: $R_3 \rightarrow G_2$. Then we move up one level, applying a_1 and a_2 on $P_{A_1}^{R_1-R_1}\ominus_{\{G_3\}}$, we can deduce from $P_{A_1}^{R_1-R_1}\ominus_{\{G_3\}}$ that the best action for R_1 is a_3: $R_1 \rightarrow S$. Similarly, we move up one more level and apply a_2, which is from $P_{A_1}^{R_1-R_3}\ominus_{\{G_3\}}$, and a_3, which is from $P_{A_1}^{R_1-R_1}\ominus_{\{G_3\}}$, to $P_{A_1}^{R_1-R_2}\ominus_{\{G_3\}}$, we can obtain a_4. Then we apply a_3, which is from $P_{A_1}^{R_1-R_1}\ominus_{\{G_3\}}$, and a_4, which is from $P_{A_1}^{R_1-R_2}\ominus_{\{G_3\}}$, to $P_{A_1}^{R_1-R_3}\ominus_{\{G_3\}}$, we can obtain a_5, and so on. After two iterations, we can see that the solution set [a_3, a_4, a_5] equals [a_6, a_7, a_8], which means the solution set no longer change. From this sub-tree, we can deduce that from R_1's point of view, R_2 will choose to obtain G_1 and R_2 will model that R_1 will stay still and R_3 will choose to obtain G_2. Similarly, from the right sub-tree of Fig. 2.12, we can deduce that from R_1's point of view, R_3 will choose to obtain G_2. So, from $P_{A_1}^{R_1-R_1}$, R_1 will choose to obtain G_3.

2.3.5 Recursive Formulas

Now, we are going to formalize the above process into *Recursive Formulas.* In the above process, excepts the lowest level, the set of best actions at each level are deduced given the set of best actions from one lower level. This relationship between the best actions is shown in Fig. 2.14. From the payoff matrix at the top level, $P_{A_1}^{R_1-R_1}$, we can deduce R_1's best action, denoted as $a_{R_1}^{*1}$, given, from R_1's point of view, the best actions of R_2 and R_3 from the second level, denoted as $a_{R_1-R_2}^{*2}$, and $a_{R_1-R_3}^{*2}$ respectively. $a_{R_1-R_2}^{*2}$, and $a_{R_1-R_3}^{*2}$ are deduced from the payoff matrices at the second level: $P_{A_1}^{R_1-R_2}\ominus_{\{G_3\}}$ and $P_{A_1}^{R_1-R_3}\ominus_{\{G_1\}}$, respectively. These two actions at the second level are deduced given the set of best actions from the third level. To deduce the set of best actions at the third level, actually we need the set of best actions from the forth level. From Fig. 2.12, we can see that in order to deduce the best action from $P_{A_1}^{R_1-R_1}\ominus_{\{G_3\}}$ at level three, we need the best actions from $P_{A_1}^{R_1-R_2}\ominus_{\{G_3\}}$ and $P_{A_1}^{R_1-R_3}\ominus_{\{G_3\}}$ from level four. Note that these two matrices actually repeat at level two and level three, respectively. So, in order to deduce the best action $a_{R_2-R_1}^{*3}$ at level three, we apply the best action from level 2: $a_{R_1-R_2}^{*2}$, deduced from $P_{A_1}^{R_1-R_2}\ominus_{\{G_3\}}$, and the best action from level 3: $a_{R_2-R_3}^{*3}$, deduced from $P_{A_1}^{R_1-R_3}\ominus_{\{G_3\}}$, to $P_{A_1}^{R_1-R_1}\ominus_{\{G_3\}}$.

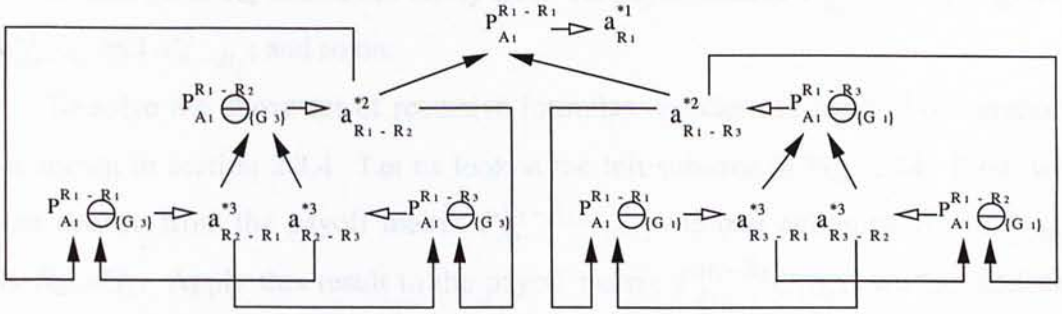

Figure 2.14: Hierarchy of best actions

From Fig. 2.14, we can write down a set of recursive formulas:

$$
\begin{cases}
1.\ a^{*1}_{R_1} = \arg\max_a U_{R_1}(P^{R_1-R_1}_{A_1}, a, a^{*2}_{R_1-R_2}, a^{*2}_{R_1-R_3}) \\
2.\ a^{*2}_{R_1-R_2} = \arg\max_a U_{R_2}(P^{R_1-R_2}_{A_1} \ominus_{\{G_3\}}, a, a^{*3}_{R_2-R_1}, a^{*3}_{R_2-R_3}) \\
3.\ a^{*3}_{R_2-R_1} = \arg\max_a U_{R_1}(P^{R_1-R_1}_{A_1} \ominus_{\{G_3\}}, a, a^{*2}_{R_1-R_2}, a^{*3}_{R_2-R_3}) \\
4.\ a^{*3}_{R_2-R_3} = \arg\max_a U_{R_3}(P^{R_1-R_3}_{A_1} \ominus_{\{G_3\}}, a, a^{*2}_{R_1-R_2}, a^{*3}_{R_2-R_1}) \\
5.\ a^{*2}_{R_1-R_3} = \arg\max_a U_{R_3}(P^{R_1-R_3}_{A_1} \ominus_{\{G_1\}}, a, a^{*3}_{R_3-R_1}, a^{*3}_{R_3-R_2}) \\
6.\ a^{*3}_{R_3-R_1} = \arg\max_a U_{R_1}(P^{R_1-R_1}_{A_1} \ominus_{\{G_1\}}, a, a^{*2}_{R_1-R_3}, a^{*3}_{R_3-R_2}) \\
7.\ a^{*3}_{R_3-R_2} = \arg\max_a U_{R_2}(P^{R_1-R_2}_{A_1} \ominus_{\{G_1\}}, a, a^{*2}_{R_1-R_3}, a^{*3}_{R_3-R_1})
\end{cases}
$$

where $a^{*1}_{R_1}$ is the best action of R_1 at level 1, which is the action that gives R_1 maximum utility from the payoff matrix $P^{R_1-R_1}_{A_1}$, given $a^{*2}_{R_1-R_2}$ and $a^{*2}_{R_1-R_3}$; and $a^{*2}_{R_1-R_2}$ is the best action of R_2 at level two, from R_1's point of view, which is the action that gives R_2 maximum utility from the payoff matrix $P_{A_1}^{R_1-R_2}\ominus_{\{G_3\}}$, given $a^{*3}_{R_2-R_1}$ and $a^{*3}_{R_2-R_3}$; and so on.

To solve the above set of recursive formulas, we can use method of iteration as shown in section 2.3.4. Let us look at the left sub-tree in Fig. 2.14. First, we can deduce from the payoff matrix $P_{A_1}^{R_1-R_2}\ominus_{\{G_3\}}$ the best action of R_2: $a^{*2}_{R_1-R_2}$ is $R_2 \rightarrow G_1$. Apply this result to the payoff matrix $P_{A_1}^{R_1-R_3}\ominus_{\{G_3\}}$, we can deduce $a^{*3}_{R_2-R_3}$, which is $R_3 \rightarrow G_2$. Then from formula 3, we can find $a^{*3}_{R_2-R_1}$, which is $R_1 \rightarrow S$. From formula 2, we can find $a^{*2}_{R_1-R_2}$ is $R_2 \rightarrow G_1$. From formula 4, we can find $a^{*3}_{R_2-R_3}$ is $R_3 \rightarrow G_2$. As the solutions do not change anymore, this iteration stops, and $a^{*2}_{R_1-R_2}$, $a^{*3}_{R_2-R_1}$, as well as $a^{*3}_{R_2-R_3}$ are found. Similarly, we can find $a^{*2}_{R_1-R_3}$ is $R_3 \rightarrow G_2$, $a^{*3}_{R_3-R_1}$ is $R_1 \rightarrow S$, and $a^{*3}_{R_3-R_2}$ is $R_2 \rightarrow G_3$. Then from formula 1, we can find $a^{*1}_{R_1}$, which is $R_1 \rightarrow G_3$. The above process means that at the third level, R_1 models how R_2 models R_1's and R_3's decision-making, as well as how R_3 models R_1's and R_2's decision-making. At the second level, R_1 models R_2's and R_3's decision-making. Since at the third level, R_1 can model that R_2 calculates that R_1 will choose to stay still and R_3 will choose to obtain G_2, R_1 can model at level two that R_2 will choose to obtain G_1. Similarly, since at level three, R_1 can model that R_3 calculates that R_1 will choose to stay still and R_2 will choose to obtain G_3, R_1 can model at level two that R_3 will choose to obtain G_2. If R_2 chooses to obtain G_1 and R_3 chooses to obtain G_2, R_1 can determine that it can get the highest payoff by obtaining G_3.

In an *n*-agent environment, we denote the payoff matrix that R_i models R_j, with level-*m* knowledge, as $P_m^{R_i-R_j}$ and we denote the best action of R_j at level-*m* from R_i's point of view as $a_m^{R_i-R_j}$. The hierarchy of best actions in From Fig. 2.15 can be formed. In the figure, except for the lowest level, the set of best actions at each level are determined given the set of best actions from one lower level. The infinite recursive hierarchy terminates at level *k* if level-(*k-1*) knowledge is not available. At the lowest level, the set of best actions are determined with the set of best action at the same level and one of the best actions from the previous level.

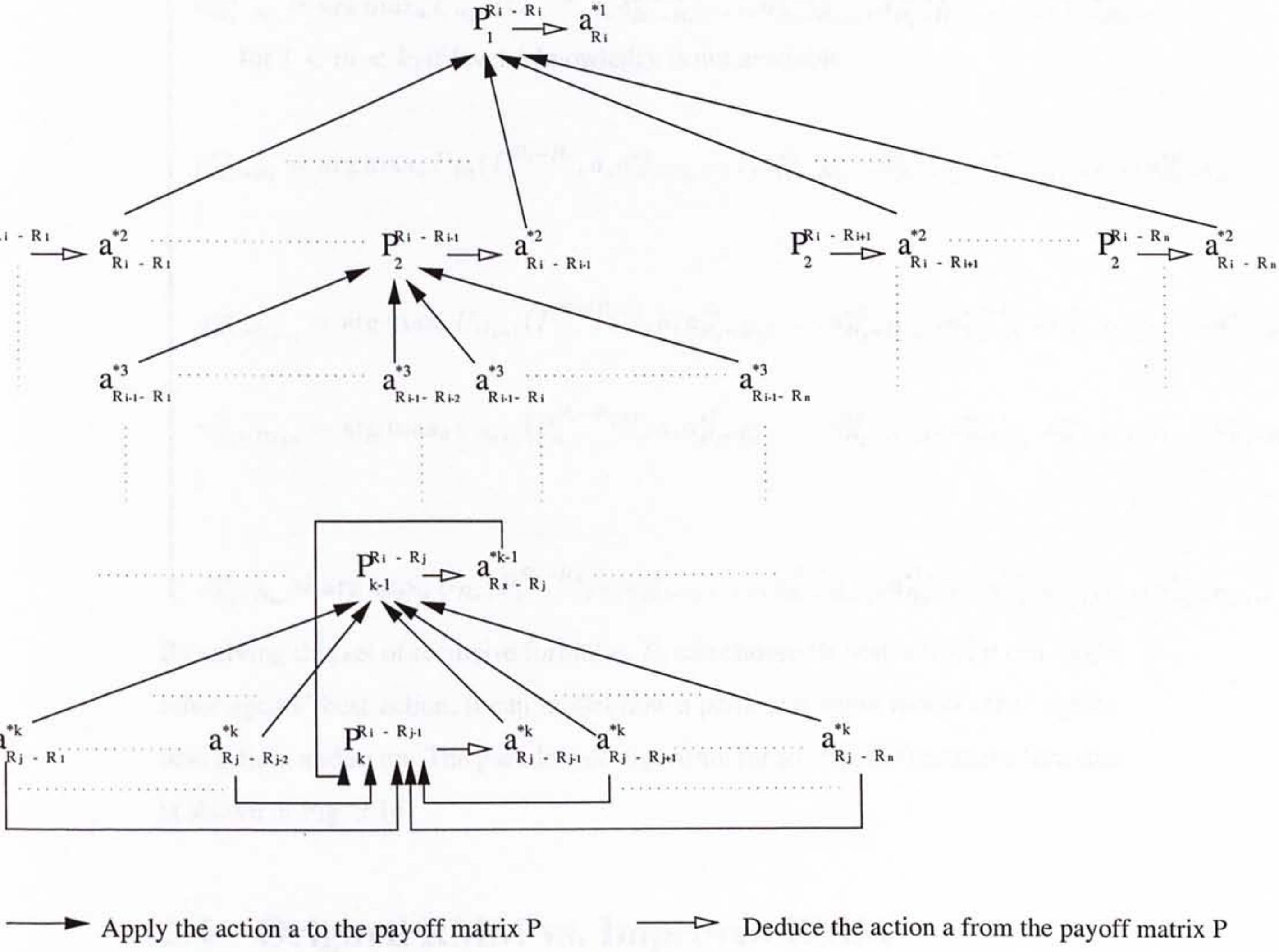

Figure 2.15: R_i's Hierarchy of best actions in an n-agent environment

From Fig. 2.15, RMM can be formulated by the following set of recursive formulas:

$$
\begin{cases}
a^{*1}_{R_i} = \arg\max_a U_{R_i}(P_1^{R_i-R_i}, a, a^{*2}_{R_i-R_1}, \ldots, a^{*2}_{R_i-R_{i-1}}, a^{*2}_{R_i-R_{i+1}}, \ldots, a^{*2}_{R_i-R_n}) \\
\\
a^{*m}_{R_x-R_j} = \arg\max_a U_{R_j}(P_m^{R_i-R_j}, a, a^{*m+1}_{R_j-R_1}, \ldots, a^{*m+1}_{R_j-R_{j-1}}, a^{*m+1}_{R_j-R_{j+1}}, \ldots, a^{*m+1}_{R_j-R_n}), \\
\quad \text{for } 1 < m < k, \text{ if level-}k\text{ knowledge is not available} \\
\\
a^{*k}_{R_j-R_1} = \arg\max_a U_{R_1}(P_k^{R_i-R_1}, a, a^{*k}_{R_j-R_2}, \ldots, a^{*k}_{R_j-R_{j-1}}, a^{*k-1}_{R_x-R_j}, a^{*k}_{R_j-R_{j+1}}, \ldots, a^{*k}_{R_j-R_n}) \\
\vdots \\
\\
a^{*k}_{R_j-R_{j-1}} = \arg\max_a U_{R_{j-1}}(P_k^{R_i-R_{j-1}}, a, a^{*k}_{R_j-R_1}, \ldots, a^{*k}_{R_j-R_{j-2}}, a^{*k-1}_{R_x-R_j}, a^{*k}_{R_j-R_{j+1}}, \ldots, a^{*k}_{R_j-R_n} \\
\\
a^{*k}_{R_j-R_{j+1}} = \arg\max_a U_{R_{j+1}}(P_k^{R_i-R_{j+1}}, a, a^{*k}_{R_j-R_1}, \ldots, a^{*k}_{R_j-R_{j-1}}, a^{*k-1}_{R_x-R_j}, a^{*k}_{R_j-R_{j+2}}, \ldots, a^{*k}_{R_j-R_n} \\
\vdots \\
\\
a^{*k}_{R_j-R_n} = \arg\max_a U_{R_n}(P_k^{R_i-R_n}, a, a^{*k}_{R_j-R_1}, \ldots, a^{*k}_{R_j-R_{j-1}}, a^{*k-1}_{R_x-R_j}, a^{*k}_{R_j-R_{j+1}}, \ldots, a^{*k}_{R_j-R_{n-1}})
\end{cases}
$$

By solving this set of recursive formulas, R_i can choose its best action, it can model other agents' best action, it can model how a particular agent model other agents' best action, and so on. The pseudocode algorithm for solving the recursive formulas is shown in Fig. 2.16.

2.4 Original RMM vs. Improved RMM

2.4.1 Terminating the Infinite Hierarchy

In the original design of RMM, the authors made an assumption that each possible action has an equal probability to be chosen. The reason to make this assumption is to terminate an infinite recursive hierarchy to a finite one when there is insufficient information. The problem is that there are errors between this approximation and

```
For v:=1, v ≦ k, v++
    For r:=1, r ≦ n, r++
        Generate payoff matrix P_v^{R_i-R_r} ;

Repeat
    Obtain best actions from payoff matrices at level-k and level-(k-1);
    Apply the best actions to payoff matrices at level-k;
Until    solution set is stable

For v:=k-1, v>1, v--
    Apply best actions at level-v to payoff matrices at level-(v-1);
    Obtain best actions at level-(v-1);
```

Figure 2.16: Pesudocode algorithm for solving recursive formulas

the real case. To solve the problem, we improve the original design with recursive formulas. In the improved design, an infinite recursive hierarchy is also terminated to a finite one when there is no further information. In this case, payoff matrices at the lowest level repeat those at upper levels. By referencing the repeated payoff matrices at the upper levels, we form the recursive formulas. By solving the recursive formulas, with no assumption made, agents can determine their best actions.

2.4.2 Resultant Payoff

Using the original RMM, with an assumed probabilistic approximation, we show in section 2.2.3 that R_1 can only choose to stay still in the example setting, with zero payoff. In section 2.3.4, we show that R_1 can determine that it can choose to obtain G_3 by solving the recursive formulas, getting a payoff of 5. In this way, R_1 can increase its utility compared to making decision with the original RMM.

2.5 Summary

Agents can use the Recursive Modeling Method (RMM) to represent their knowledge, predict other agents' actions and then choose their own actions, that is to choose which goal to obtain or choose whether to stay still. In this chapter, we present and compare the original design and the modified design of RMM by an illustrative example. The modified design of RMM improves the probabilistic approximation of the original one by recursive formulas.

In this thesis, we have not mention how the recursive formulas can be solved in a computationally tractable way, which is out of the scope of this thesis. As future work, we are going to design an algorithm for solving the recursive formulas. In addition, we are going to compare the performance of the original RMM with that of the improved RMM through simulations.

Chapter 3

A Trust/Honesty Model

RMM describes a decision making strategy for agents to choose which goal to obtain. In a semi-competitive environment, an agent has motivation to tell the truth when it wants to invite another agent to cooperate. On the other hand, an agent has motivation to tell lie when it wants to mislead other agents. As a result, on receiving a message, the receiver needs to decide whether to believe the message or not. To help receivers choose such actions, we propose a Trust Model. As receivers employ a Trust Model, receivers become less easy to cheat. As a result, senders cannot always tell lies. To help senders determine whether to tell lies, we propose an Honesty Model. In this chapter, we present the Trust/Honesty Model, which helps agents choose such actions. We first introduce the needs for the model, and then we present the model.

3.1 The Need for a Trust Model

The following example shows that agents have motivations to tell the truth and agents also have motivations to tell lies in a semi-competitive environment. As a result, receivers need to choose whether to believe the message or which message to believe. The example also shows that agents cannot make the decision by considering only the expected payoffs of the messages. To help agents make the decisions, we need a Trust Model.

3.1.1 Motivation to Tell the Truth: Invitation to Cooperate

Agents are self-interested, and always want to maximize their own respective utilities. Therefore, in the semi-competitive environment, it is not always good for an agent to share all of its information with other agents. Consider the example described in section 2.1. From section 2.3.5, by recursive formulas, R_1 chooses to obtain G_3, getting a payoff of 5. In addition, it models that R_2 will choose to obtain G_1 and R_3 will choose to obtain G_2. However, from R_1's payoff matrix in Fig. 2.3, R_1 knows that it can increase its payoff by obtaining G_1, getting a payoff of 8, if R_2 chooses to obtain G_3 instead of G_1. To invite R_2 for cooperation, R_1 considers it rational to send R_2 the information about G_3, which is unknown to R_2. This message, M_1, should look like this: "You can obtain the goal G_3, with worth 20, *cost*($R_1 \rightarrow G_3$) = 15, *cost*($R_2 \rightarrow G_3$) = 2, *cost*($R_3 \rightarrow G_3$) = 5 and G_3's priority list is $< R_2, R_3, R_1 >$."

After communication, if R_2 believes the message M_1, the model of R_2's decision-making situation will be changed. The payoff matrix describing R_2's decision-making situation with the new knowledge of the presence of G_3 will become the one in Fig. 2.9. From which, it can be seen that if R_2 believes the message M_1, it will choose to obtain G_3. This is because R_2 can get a payoff of 18 by obtaining G_3, no matter what actions other agents take. In fact this payoff is also the best payoff it can get among all its possible actions. If R_2 believes and follows the message M_1, R_1 can also increase its payoff from 5 to 8 by obtaining G_1 instead of obtaining G_3. This shows an example in which agents can benefit mutually by cooperation. This also shows that agents have incentives to tell the truth.

3.1.2 Motivation to Tell a Lie: to Prevent Competition

At the same time, R_3 can maximize its payoff by obtaining G_3. However, it has a lower priority than R_2, which means that it needs to compete with R_2. So, to prevent competition with R_2, R_3 considers it rational lying to R_2 and directing it to

		Nature	
		M_1 is True	M_1 is False
R_2	*Believe*	18	−2
	Not Believe	6	6

Table 3.1: The payoffs of R_2 with respect to its trust on M_1 and the nature of M_1

		Nature	
		M_2 is True	M_2 is False
R_2	*Believe*	20	−4
	Not Believe	6	6

Table 3.2: The payoffs of R_2 with respect to its trust on M_2 and the nature of M_2

a fake goal. This message, M_2, should look like this: "You can obtain the goal G_4, with worth 24, $cost(R_1 \rightarrow G_4) = 50$, $cost(R_2 \rightarrow G_4) = 4$, $cost(R_3 \rightarrow G_4) = 50$ and G_4's priority list is $< R_2, R_1, R_3 >$." This shows an incentive for an agent to lie.

3.1.3 To Believe, or Not to Believe, that is the Question

Now R_2 receives two messages: M_1 from R_1 and M_2 from R_3. If R_2 believe M_1, it will choose to obtain G_3. In this way, R_2 can gain a payoff of 18 if the message M_1 is true, and loss the cost of 2 if the message M_1 is a lie. On the other hand, if R_2 does not believe the message M_1, it will choose to obtain G_1 and get a payoff of 6 no matter the message M_1 is true or not. The payoffs of R_2 with respect to its trust on M_1 as well as the nature of M_1 are summarized in Table 3.1. However, if R_2 believes M_2, it will choose to obtain G_4. In this way, R_2 can gain a payoff of 20 if the message M_2 is true, and loss the cost of 4 if the message M_2 is a lie. Otherwise, it will choose to obtain G_1, getting a payoff of 6 no matter the message M_2 is true or not. The payoffs of R_2 with respect to its trust on M_2 as well as the nature of M_2 are summarized in Table 3.2.

Now, R_2 faces a difficult question. If R_2 makes the simple assumption that the

	Expected Utility
Believe and follow M_1	$\frac{1}{2} \times (18-2)$=8
Believe and follow M_2	$\frac{1}{2} \times (20-4)$=8
Not Believe M_1	$\frac{1}{2} \times (6+6)$=6
Not Believe M_2	$\frac{1}{2} \times (6+6)$=6

Table 3.3: Resulting expected utilities

probability for R_1 or R_3 telling the truth to be $\frac{1}{2}$, then the resulting expected utilities are shown in Table 3.3. From the table, it can be seen that both the expected utilities of believing and following M_1 as well as M_2 are higher than that of believing neither M_1 nor M_2, so R_2 will believe either M_1 or M_2, or both. However, believing M_1 will lead R_2 to obtain the goal G_3 and believing M_2 will lead R_2 to obtain the goal G_4, which are two different actions. Since each agent can only take one action, R_2 has to choose to follow either M_1 or M_2, but not both. As the expected utilities of believing and following M_1 and believing and following M_2 are the same, R_2 cannot determine whether to follow M_1 or M_2. [1]

3.2 The Trust Model

3.2.1 Impression

From Cambridge Dictionaries Online [htta], *Impression* is "the opinion you form when you meet someone or see something."

From Merriam-Webster Online [httb], *Impression* is "a telling image impressed on the senses or the mind."

We suggest that in semi-competitive environments, each receiver should maintain an impression on each sender based on its experience. A sender gives a good

[1] In this paper, we assume that if an agent believes a message, the agent believes all the information provided by the message. It is arguable that an agent can, in general, choose to believe only some parts of the message. However, this is not our scope of discussion.

impression to a receiver if and only if the former has told truths to the latter, which has brought the latter benefits. Follow the definitions in the dictionaries [htta, httb], we define the *impression* that receiver *i* has towards sender *j* to be a real number in $[-1,1]$:

$$imp_{ij} = f_i(\sum gain_{ij}, \sum loss_{ij}, p, n)$$

where $\sum gain_{ij}$ is the sum of the utility that agent *i* has gained by having believed the truths from agent *j*, $\sum loss_{ij}$ is the sum of the utility that agent *i* has lost by having believed the lies from agent *j*, *p* is the number of times that agent *j* has told the truth, and *n* is the total number of messages that agent *i* has received from agent *j*. The function f_i must satisfy the following axioms:

Axiom f_{i1}: f_i is continuous.

Axiom f_{i2}: f_i strictly increases as *p* increases.

Axiom f_{i3}: f_i increases as $\sum gain_{ij}$ increases.

Axiom f_{i4}: f_i decreases as $\sum loss_{ij}$ increases.

Axiom f_{i5}: $f_i = 0$ when $n = 0$.

Axiom f_{i6}: For $\sum gain_{ij} = \sum loss_{ij}$, $f_i = 0$ when $p = n - p$, $f_i > 0$ when $p > n - p$, and $f_i < 0$ when $p < n - p$.

Axiom f_{i7}: $f_i > 0$ when $\sum gain_{ij} > \sum loss_{ij}$ and $p \geq n - p$.

Axiom f_{i8}: $f_i < 0$ when $\sum gain_{ij} < \sum loss_{ij}$ and $p \leq n - p$.

Axiom f_{i9}: $f_i < 0$ when $\sum gain_{ij} > \sum loss_{ij}$ and $p < n - p$.

Axiom f_{i10}: $f_i < 0$ when $\sum gain_{ij} < \sum loss_{ij}$ and $p > n - p$.

Axiom f_{i2} states that it is rational that impression will increase if the number of times that the message sender has told the truth to the receiver increases. Axiom f_{i3} means it is rational that impression will increase if the sum of the utility that the receiver has gained by having believed the messages from the sender increases. Axiom f_{i4} describes that it is also rational that impression will decrease if the sum of the utility that the receiver has lost by having believed the messages from the sender increases.

Axiom f_{i5} says that impression will be neutral if agent *i* receives no message from agent j. For the gain in utility equals the loss in utility (axiom f_{i6}), the impression will also be neutral if the message sender has told the same number of truths and lies, the impression will be positive if the sender has told more truths than lies, and the impression will be negative if the sender has told more lies than truths.

From axiom f_{i7}, if the gain in utility is greater than the loss in utility and the sender has told more truths than lies (or the same number of truths and lies), this means that the message sender is good to the receiver, so the receiver will have a positive impression towards the message sender. On the other hand, from axiom f_{i8}, if the loss in utility is greater than the gain in utility and the sender has told more lies than truths (or the same number of truths and lies), this means that the message sender is doing harm to the receiver, so the receiver will have a negative impression.

Axiom f_{i9} is special. If the gain in utility is greater than the loss in utility, but the sender has told more lies than truths, it is very likely that the sender is performing some kinds of strategy. For example, at the first encounter, the sender tells a truth, bringing a utility of 100 to the receiver; but in the following nine encounters, the sender lies, which makes the receiver loss a utility of 90 in total, it is obvious that the sender is doing harm to the receiver. So, the impression in this case should be negative. Axiom f_{i10} shows the case in which the sender has told more truths than lies, but the gain in utility is less than loss in utility, this means that the lies bring more harms to the receiver, so the impression is also negative.

The following is an example function satisfying the above axioms and the intuitive meanings:

$$imp_{ij} = \begin{cases} 0 & n = 0 \\ \frac{p-(n-p)}{n} & \sum gain_{ij} = \sum loss_{ij} \\ -(\frac{p-(n-p)}{n})(\frac{\sum gain_{ij} - \sum loss_{ij}}{\sum gain_{ij} + \sum loss_{ij}}) & \sum gain_{ij} < \sum loss_{ij} \bigwedge p \leq n-p \\ (\frac{p-(n-p)}{n})(\frac{\sum gain_{ij} - \sum loss_{ij}}{\sum gain_{ij} + \sum loss_{ij}}) & \text{otherwise} \end{cases}$$

3.2.2 Reputation

From Cambridge Dictionaries Online [htta], *Reputation* is "the opinion that people in general have about someone or something, ... , based on past behavior or character."

From Merriam-Webster Online [httb], *Reputation* is the "overall quality or character as seen or judged by people in general."

Follow the definitions in the dictionaries [htta, httb], we define the reputation of an agent to be an averaged impression that the population has towards that agent. However, the only way for an agent to access other agents' impressions on a particular agent is to ask other agents for their impressions on that particular agent. It is possible that an agent can lie in answering the query, so a weight could be introduced to the answer. In an N agents environment, we define *reputation* of a sender j, as seen by a receiver i, as a weighted sum of individual impressions of a subset of the population:

$$rep_{ij} = \frac{\sum_{k=1}^{k=n} imp_{kj} \times W_{ik}}{n}$$

where W_{ik} is the weight that agent i attaches to agent k's impression on agent j and $n \leq N$. We shall sometimes omit the phrase "as seen by agent i" when the meaning is unambiguous from the context. Note that each receiver can choose its own subset of population and decide the corresponding weights in calculating the reputation of a particular sender. Much research has been done on this issue [MMH02, RLM01, SS01]. In the absence of any knowledge about other agents' honesty and trustworthiness, the weights can be assumed to be 1.

3.2.3 Risk Attitude and Trustworthiness

In human interaction, different people have different reactions when they are cheated by the same lie, and the degree of trust that different people have towards the liar will be different. For example, one will consider not trusting the liar anymore once

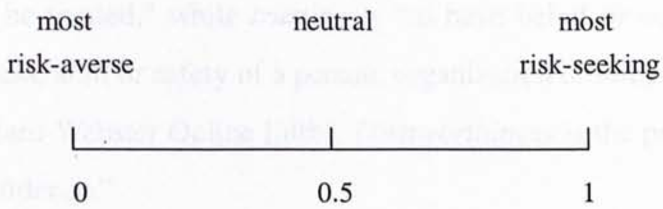

Figure 3.1: Risk attitude

he is cheated, while another person may continue trusting the liar even he is cheated. This is because different people have different *attitudes towards risk*: some do not mind taking any risk, some do not want to take any risk, while others are neutral. To model this, we propose to include the *risk attitude* of the receiver in calculating its trustworthiness on the sender. The risk attitude here does not mean the risk undertaken by the agent, but rather an index, which reflects the amount of risk that the agent is willing to undertake. Here, we define *risk attitude*, *r*, of an agent to be a real number in [0,1], which is shown in Fig. 3.1. Agents with risk attitude being 0 are the most risk-averse while agents with risk attitude being 1 are the most risk-seeking. A risk-averse agent prefers messages from a sender with high trustworthiness, while a risk-seeking agent prefers messages with high utilities. This risk attitude is determined by the agent itself, like the personality of human, and can change over time.

In the example shown in section 3.1.3, it can be seen that considering expected utility alone is not enough for an agent to determine which message(s) to believe and follow when multiple messages are received. In fact, it is dangerous for an agent to believe and follow a message just because the expected utility of the message is attractive: the agent can be cheated easily. We propose that in a multiagent semi-competitive environment, each receiver should maintain a *trustworthiness* to every sender of the messages it receives. In other words, for each ordered pair $< R_1, R_2 >$ we associate a *trustworthiness of* R_2 *as seen by* R_1. We shall sometimes omit the phrase "as seen by R_1" when the meaning is unambiguous from the context.

From Cambridge Dictionaries Online [htta], *Trustworthiness* is the property of being "able to be trusted," while *trusting* is "to have belief or confidence in the honesty, goodness, skill or safety of a person, organization or thing."

From Merriam-Webster Online [httb], *Trustworthiness* is the property of being "worthy of confidence."

We define *trustworthiness*, t_{ij}, that receiver *i* has towards sender *j* as a function of the impression that agent *i* has about agent *j*, agent *i*'s calculation on the reputation of agent *j* as well as the risk attitude of agent *i*:

$$t_{ij} = f_t(imp_{ij}, rep_{ij}, r_i)$$

The function f_t returns a real number in $[-1, 1]$, and must satisfy the following axioms:

Axiom f_{t1}: f_t is continuous.

Axiom f_{t2}: f_t decreases as imp_{ij} decreases and vice versa.

Axiom f_{t3}: f_t decreases as rep_{ij} decreases and vice versa.

Axiom f_{t4}: f_t decreases as r_i decreases and vice versa.

Axiom f_{t2} states that it is rational that the trustworthiness of the sender decreases if the receiver's impression on it decreases and vice versa. Similarly, axiom f_{t3} states that it is rational that the trustworthiness of the sender decreases if its reputation decreases and vice versa. Axiom f_{t4} states that if the risk attitude of the receiver decreases, which means the receiver becomes more risk-averse and thus less willing to trust other agents, the evaluated trustworthiness of the sender will decreases.

An example function satisfying the above axioms and the intuitive meanings is shown below, which attaches the same degree of importance to impression and reputation, and is in proportion to the agent's risk attitude.

$$t_{ij} = \frac{imp_{ij} + rep_{ij}}{2} \times (1 - r_i)$$

3.2.4 Persuasiveness of a Message vs. Stubbornness of the Receiver

On determining whether to believe and follow a particular message, besides considering the expected payoffs that the receiver can gain by believing the message, trustworthiness of the message sender should also be considered. Formally, a receiver makes use of a *persuasiveness function* f_p to rank the messages and choose to follow the message that has the highest value of persuasiveness. The *persuasiveness*, p_M, of a message M is defined by:

$$p_M = f_p(r_i, t_{ij}, u_k)$$

Intuitively, the function f_p takes the risk attitude r_i of the receiver i as the first argument, the trustworthiness t_{ij} of the message sender j, as seen by receiver i, as the second argument and the expected utility u_k of the message k as the last argument, and returns a real number in $[-1, 1]$ as the rank of the message. The function f_p must satisfy the following axioms:

Axiom f_{p1}: f_p is continuous.

Axiom f_{p2}: (*Adventurousness of risk-seeking agents*) There exists a value $r_0 \in \Re$ such that $f_p(r, t_2, u_1) > f_p(r, t_1, u_2)$ if and only if $r > r_0$, $t_1 > t_2$ and $u_1 > u_2$.

Axiom f_{p3}: (*Cautiousness of risk-averse agents*) There exists a value $r'_0 \in \Re$ such that $f_p(r, t_2, u_1) < f_p(r, t_1, u_2)$ if and only if $r < r'_0$, $t_1 > t_2$ and $u_1 > u_2$.

Axiom f_{p4}: if $u_1 \geq u_2$, $f_p(r, t, u_1) \geq f_p(r, t, u_2)$ for $r > r_0$ and $f_p(r, t, u_1) \leq f_p(r, t, u_2)$ for $r < r'_0$.

Axiom f_{p5}: if $t_1 \geq t_2$, $f_p(r, t_1, u) \geq f_p(r, t_2, u)$.

Axiom f_{p6}: if $r_1 \geq r_2$, $f_p(r_1, t, u) \geq f_p(r_2, t, u)$.

It is obvious that the domains of the inputs of f_p are continuous, so f_p should be continuous. Besides, it is reasonable that utility will be more attractive than the trustworthiness of the message sender to a risk-seeking receiver, and vice versa to a risk-averse receiver. These bring about axiom f_{p2} and Axiom f_{p3}. Axiom f_{p4} states that if a receiver receives two messages from senders with the same trustworthiness, but with different payoffs, it is rational that a risk-seeking receiver is more willing to follow the message with higher payoff. However, it is rational that a risk-averse receiver is cautious for the message with higher payoff. Axiom f_{p5} states that if a receiver receives two messages from senders with different trustworthiness, but with same payoffs, it is rational that the receiver is more willing to follow the message from sender, which is more trustworthy. Axiom f_{p6} means that the persuasiveness of the same message from the same sender, with the same trustworthiness, decreases if the receiver become more risk-averse.

Theorem 1. *r_0 (in axiom f_{p2}) equals r'_0 (in axiom f_{p3}).*

Proof. Assume r_0 is not equal to r'_0, this results in the following two cases:

Case 1. $r'_0 > r_0$ By axiom f_{p2}, if $r > r_0$, $t_1 > t_2$ and $u_1 > u_2$, $f_p(r, t_2, u_1) > f_p(r, t_1, u_2)$. By axiom f_{p3}, if $r < r'_0$, $t_1 > t_2$ and $u_1 > u_2$, $f_p(r, t_2, u_1) < f_p(r, t_1, u_2)$. As a result, for $r_0 < r < r'_0$, $f_p(r, t_2, u_1) > f_p(r, t_1, u_2)$ and $f_p(r, t_2, u_1) < f_p(r, t_1, u_2)$, which is a contradiction.

Case 2. $r'_0 < r_0$ By axiom f_{p2}, $f_p(r, t_2, u_1) > f_p(r, t_1, u_2)$ if and only if $r > r_0$, $t_1 > t_2$ and $u_1 > u_2$, so for $r < r_0$, $f_p(r, t_2, u_1) \leq f_p(r, t_1, u_2)$. By axiom f_{p3}, $f_p(r, t_2, u_1) < f_p(r, t_1, u_2)$ if and only if $r < r'_0$, $t_1 > t_2$ and $u_1 > u_2$, so for $r > r'_0$, $f_p(r, t_2, u_1) \geq f_p(r, t_1, u_2)$. As a result, for $r'_0 < r < r_0$, $f_p(r, t_2, u_1) \leq f_p(r, t_1, u_2)$ and $f_p(r, t_2, u_1) \geq f_p(r, t_1, u_2)$, which is a contradiction.

So r_0 equals r'_0. □

A simple example satisfying the above axioms, f_p can be defined as follows:[2]

$$f_p(r,t,u) = \begin{cases} \frac{(r-1)u+t}{2} & \text{for } r < 0.5 \\ \frac{u+t}{2} & \text{for } r = 0.5 \\ \frac{(r+1)u+t}{3} & \text{for } r < 0.5 \end{cases}$$

[2]In this formula, the utility is assumed to be in the range $[0, 1]$.

With this function, the more risk-averse the receiver is, the more important the trustworthiness of the sender is in making decision. The more risk-seeking the receiver is, the more important the utility of the message is, while receivers with neutral risk attitude consider trustworthiness of the sender and utility of the message to be the same important.

If a receiver receives only one message, it should believe and follow the message only if the persuasiveness of the message is higher than a certain threshold. We call this the *stubbornness* of the receiver to the sender, which is a real number in $[-1,1]$. We shall sometimes omit the phrase "to the sender" when the meaning is unambiguous from the context. Each receiver maintains a stubbornness to each message sender, which can be changed over time, like personality of human. If more than one message is received at a time, as an agent can only choose one action in one single round, the receiver should believe and follow the message with the greatest persuasiveness, among those messages having a persuasiveness greater than the corresponding stubbornness to the senders.

From the definition of the f_p function, it is easy to see that it is possible that two messages have the same value of persuasiveness. This means that the two messages apparently are having the same expected utility and both are from sources with the same degree of reliability. In this case, the effect on believing and following which message will have no difference, so the agent can simply throw a dice to determine which message to believe and follow. Another problem is that a message with an extremely high utility will cause a risk-seeking agent to follow. First, we note that this actually mimics a real-life phenomenon occurring in human community. Second, at the end of a round when the worth of the goals are given to the agents, an agent actually will know whether it has believed and followed a true message or a lie, and a cheated agent then decrease its impression on the message sender who lied to it, and thus decrease the trustworthiness of the liar. In *Iterated Game* described in the following section, the impression, reputation, trustworthiness, risk attitude, and stubbornness of agents preserve in the transition from one round to another round, an agent will be cheated for only the first few times, and will not believe further messages from the same message sender.

Theorem 2. *If two messages M_1 and M_2, with expected utilities u_1, u_2 and trustworthiness of message senders t_1, t_2, respectively, where $u_1 > u_2$ and $t_2 > t_1$, are sent to all receivers with different risk attitudes. Then there exists a constant $r_0 \in \Re$ depending only on u_1, u_2 and t_1, t_2, such that all the receivers with risk attitude $r > r_0$ will choose to believe and follow M_1 and all the receivers with risk attitude $r < r_0$ will choose to believe and follow M_2 if persuasiveness of the messages are greater than the receivers' stubbornness.*

Proof. Receiver uses a function f_p to rank the messages and choose to believe and follow the message that has the highest value of f_p, where f_p must satisfy axioms f_{p1} to axioms f_{p6}. Since $t_2 > t_1$ and $u_1 > u_2$, by axiom f_{p2}, there exists a value $r_0 \in \Re$ such that $f_p(r, t_1, u_1) > f_p(r, t_2, u_2)$ if and only if $r > r_0$. And by axiom f_{p3}, there exists a value $r'_0 \in \Re$ such that $f_p(r, t_1, u_1) < f_p(r, t_2, u_2)$ if and only if $r < r'_0$. By Theorem 1, $r = r'_0$. So, for $u_1 > u_2$ and $t_2 > t_1$, there exists a value $r_0 \in \Re$ such that if $r > r_0$, then $f_p(r, t_1, u_1) > f_p(r, t_2, u_2)$, which means the receiver will choose to believe and follow message M_1, and if $r < r_0$, then $f_p(r, t_1, u_1) < f_p(r, t_2, u_2)$, which means the receiver will choose to believe and follow message M_2 if persuasiveness of the messages are greater than the receivers' stubbornness. □

Theorem 3. *Suppose there are two receivers R_1 and R_2, with risk attitudes r_1 and r_2 respectively, where $r_1 > r_2$, that is receiver R_1 is more risk-seeking than receiver R_2. Then there exist two messages M_1 and M_2, with expected utilities u_1, u_2 and trustworthiness of message senders t_1, t_2, respectively, where $u_1 > u_2$ and $t_2 > t_1$, such that when these two messages are sent to R_1 and R_2, R_1 will choose to believe and follow message M_1 and R_2 will choose to believe and follow message M_2 if persuasiveness of the messages are greater than the receivers' stubbornness.*

Proof. By theorem 2, for any two messages M_1 and M_2, with expected utilities u_1, u_2 and trustworthiness of message senders t_1, t_2, respectively, where $u_1 > u_2$ and

$t_2 > t_1$, there exists a constant $r_0 \in \Re$ depending only on u_1, u_2 and t_1, t_2, such that all the receivers with risk attitude $r > r_0$ will choose to believe and follow M_1 and all the receivers with risk attitude $r < r_0$ will choose to believe and follow M_2 if persuasiveness of the messages are greater than the receivers' stubbornness. In other words, proving theorem 3 is to find the two messages such that $r_2 < r_0 < r_1$. We do this by first initialize two messages M_1 and M_2, with expected utilities u_1, u_2 and trustworthiness of message senders t_1, t_2, respectively, where $u_1 > u_2$ and $t_2 > t_1$. When these two messages are sent to the two receivers, one of the following four cases will result:

Case 1. Both of R_1 and R_2 choose to believe and follow message M_1. In this case, generate another two messages M_1' and M_2', with expected utilities u_1', u_2' and trustworthiness of message senders t_1', t_2', respectively, where $u_1' > u_2'$, $t_2' > t_1'$, $t_2' > t_2$ and $t_1' < t_1$.

Case 2. Both of R_1 and R_2 choose to believe and follow message M_2. In this case, generate another two messages M_1' and M_2', with expected utilities u_1', u_2' and trustworthiness of message senders t_1', t_2', respectively, where $u_1' > u_2'$, $t_2' > t_1'$, $u_1' > u_1$ and $u_2' < u_2$.

In case 1 and case 2, the process is continued by sending the new messages M_1' and M_2' to the agents, replacing the old messages M_1 and M_2. As r is a real number, as long as $r_1 > r_2$, there exists $r_0 \in \Re$, such that $r_2 < r_0 < r_1$. So, eventually, the process converge and case 3 will results.

Case 3. R_1 chooses to believe and follow message M_1 and R_2 chooses to believe and follow message M_2. In this case, the theorem is proved.

Case 4. R_1 chooses to believe an follow message M_2 and R_2 chooses to believe and follow message M_1. In fact, this case will never happen. Suppose R_1 and R_2 choose to believe and follow different messages, as $r_1 > r_2$, and by theorem 2, $r_2 < r_0 < r_1$, which means R_1 will choose to believe and follow message M_1 and R_2 will choose to believe and follow message M_2. □

The following theorem states that it is rational for a risk-seeking receiver to believe a message with a higher utility and from a more trustworthy source, rather than a message with a lower utility and from a less trustworthy source. For a risk-averse receiver, as it will be cautious for the message with a higher utility, which message it believes depends on the actual values of the trustworthiness, utilities, risk attitude and stubbornness.

Theorem 4. *For risk attitude $r > r_0$, trustworthiness t_1 and t_2, and utilities u_1 and u_2, where $t_1 \geq t_2$ and $u_1 \geq u_2$, $f_p(r, t_1, u_1) \geq f_p(r, t_2, u_2)$.*

Proof. From axiom f_{p4}, if $u_1 \geq u_2$, $f_p(r, t_1, u_1) \geq f_p(r, t_1, u_2)$ for $r > r_0$. From axiom f_{p5}, if $t_1 \geq t_2$, $f_p(r, t_1, u_2) \geq f_p(r, t_2, u_2)$. So, $f_p(r, t_1, u_1) \geq f_p(r, t_1, u_2) \geq f_p(r, t_2, u_2)$. That is $f_p(r, t_1, u_1) \geq f_p(r, t_2, u_2)$. □

Intuitively, if a receiver becomes more risk-averse and lowers the trustworthiness of the message sender after it is being cheated, then when this receiver receives the same message from the same sender (with trustworthiness lowered), it should be less willing to follow the message. This phenomenon is confirmed by the following theorem.

Theorem 5. *For risk attitudes r_1 and r_2, trustworthiness t_1 and t_2, and utility u, where $r_1 \geq r_2$ and $t_1 \geq t_2$, $f_p(r_1, t_1, u) \geq f_p(r_2, t_2, u)$.*

Proof. From axiom f_{p6}, if $r_1 \geq r_2$, $f_p(r_1, t_1, u) \geq f_p(r_2, t_1, u)$. From axiom f_{p5}, if $t_1 \geq t_2$, $f_p(r_2, t_1, u) \geq f_p(r_2, t_2, u)$. So, $f_p(r_1, t_1, u) \geq f_p(r_2, t_1, u) \geq f_p(r_2, t_2, u)$. That is $f_p(r_1, t_1, u) \geq f_p(r_2, t_2, u)$. □

3.3 The Need for an Honesty Model

We define an *Iterated Game* to be a game consists of a series of Single-round games, in which one round of game proceeds after another. There is a completely new set of goals in each round of game. The impression, reputation, trustworthiness, risk attitude, and stubbornness of agents preserve in the transition from one round to another round and the values will be updated at the end of each round.

In a Single-round Game, receivers can only discover the truth at the end of the game. So, senders can always tell lies, because lying can bring utility gain but brings no penalty in a Single-round Game. However, in Iterated Games, after a receiver discovers that it is cheated, it will rationally decrease the sender's trustworthiness, and in addition it may become more risk-averse and stubborn, so as to prevent itself from being cheated again. As a result, although lying brings an increase in utility, lying in Iterated Games also brings a penalty of lost in trustworthiness. In this section, we show the needs for a sender to decide whether or not to tell lies and we show how it can do so in the next section.

3.3.1 To Lie, or Not to Lie, that is the Question

For naive receivers that do not employ any trust model, it is rational for a sender to lie, if it can model that the receiver will believe the message and change its action accordingly, which brings the sender an increase in utility. In fact, a sender can lie that the worth of a fake goal is extremely large, so that it can always be sure that the receiver will believe the message as receivers with no trust model consider only expected utility. This means that agents will always choose to lie. However, the receivers become less easy to be cheated after employing a trust model. In addition to the expected utility, a receiver also takes into account the trustworthiness of the message sender, when it decides whether to believe the received message. As a result, a sender also needs to consider if the receiver will actually be cheated before telling lies. So, whether or not to lie becomes a question.

3.3.2 Problem of Living a Lie

Suppose R_a knows that R_b has a higher priority than R_a in all the goals' priority ordering of agents. In order for R_a to get any worth, it must direct R_b to some fake goals. This can be done by sending a message to R_b, which looks like: "There is a goal G, which worths 1000 and only costs you 10." If R_b is risk-seeking enough, or if the worth of the goal is attractive enough, R_b will believe the message and choose to obtain the fake goal. In this way, R_a can obtain any goal it wants. However, no agents can live a lie in Iterated Games. It is because whenever an agent discovers that it is cheated by another agent, its impression on the liar, the liar's reputation, as well as the trustworthiness of the liar will be decreased. Eventually, that agent being cheated will stop believing the lair. In this example, if R_a tells a lie to R_b, which makes R_b loss utility, R_b's impression on R_a and R_a's reputation will be decreased and so as the trustworthiness of R_a. After several iterations, R_b may no longer believe R_a anymore. As a result, agents need to choose whether to tell lies or not in an Iterated Game.

3.4 The Honesty Model

3.4.1 Impression

In semi-competitive environments, each sender also maintains an impression on each receiver, based on its past experience. We define the *impression* that sender i has towards receiver j to be a real number in $[-1, 1]$:

$$imp_{ij} = f_i(\sum gain_{ij}, \sum loss_{ij}, p, n)$$

where $\sum gain_{ij}$ is the sum of the utility that agent i has gained by successfully cheating agent j, $\sum loss_{ij}$ is the sum of the utility that agent i has lost by unsuccessfully cheating agent j, p is the number of times that agent j has been successfully cheated by agent i, and n is the total number of times that agent i lie to agent j. This function follows the same set of axioms as described in section 3.2.1:

Axiom f_{i1}: f_i is continuous.

Axiom f_{i2}: f_i strictly increases as p increases.

Axiom f_{i3}: f_i increases as $\sum gain_{ij}$ increases.

Axiom f_{i4}: f_i decreases as $\sum loss_{ij}$ increases.

Axiom f_{i5}: $f_i = 0$ when $n = 0$.

Axiom f_{i6}: For $\sum gain_{ij} = \sum loss_{ij}$, $f_i = 0$ when $p = n - p$, $f_i > 0$ when $p > n - p$, and $f_i < 0$ when $p < n - p$.

Axiom f_{i7}: $f_i > 0$ when $\sum gain_{ij} > \sum loss_{ij}$ and $p \geq n - p$.

Axiom f_{i8}: $f_i < 0$ when $\sum gain_{ij} < \sum loss_{ij}$ and $p \leq n - p$.

Axiom f_{i9}: $f_i < 0$ when $\sum gain_{ij} > \sum loss_{ij}$ and $p < n - p$.

Axiom f_{i10}: $f_i < 0$ when $\sum gain_{ij} < \sum loss_{ij}$ and $p > n - p$.

Axioms f_{i2} and f_{i3} state that sender will have a better impression on the receiver if the number of times that the receiver is cheated by the sender increases, or the sum of utility that the sender has gained from the receiver increases. On the other hand, axiom f_{i4} states that impression decreases when the sum of utility that the sender has lost increases due to the receiver's distrust on it.

Axioms f_{i5} and f_{i6} say that impression will be neutral if there is no interaction between the sender and the receiver, or if the sender gains as much as loses and the receiver is cheated successfully and unsuccessfully for the same number of times, while impression will be positive if the number of times that the receiver is cheated successfully is more than that of unsuccessfully, and vice versa.

From axioms f_{i7} and f_{i8}, impression is positive if the sender gains more than loses and the number of times that the receiver is cheated successfully is more than (or equal to) that of unsuccessfully and vice versa. Axiom f_{i9} state that even if the sender gains more than loses but the number of times that the receiver is cheated successfully is less than that of unsuccessfully, the sender should be cautious for this receiver and the impression is negative. Similarly, impression should also be negative if the sender loses more than gains even if the number of times that the receiver is cheated successfully is more than that of unsuccessfully, which is axiom

f_{i10}.

The following is an example function satisfying the above axioms and the intuitive meanings:

$$imp_{ij} = \begin{cases} 0 & n = 0 \\ \frac{p-(n-p)}{n} & \sum gain_{ij} = \sum loss_{ij} \\ -(\frac{p-(n-p)}{n})(\frac{\sum gain_{ij} - \sum loss_{ij}}{\sum gain_{ij} + \sum loss_{ij}}) & \sum gain_{ij} < \sum loss_{ij} \bigwedge p \leq n-p \\ (\frac{p-(n-p)}{n})(\frac{\sum gain_{ij} - \sum loss_{ij}}{\sum gain_{ij} + \sum loss_{ij}}) & \text{otherwise} \end{cases}$$

3.4.2 Reputation

Similarly, each sender also maintains a reputation on each receiver about ease of being cheated by asking other agents for their impressions on that particular agent. It is also possible that an agent can lie in answering the query, so a weight could be introduced to the answer. In an *N* agents environment, we define *reputation* of a receiver *j*, as seen by a sender *i*, as a weighted sum of individual impressions of a subset of the population:

$$rep_{ij} = \frac{\sum_{k=1}^{k=n} imp_{kj} \times W_{ik}}{n}$$

where W_{ik} is the weight that agent *i* attaches to agent *k*'s impression on agent *j* and $n \leq N$. In the absence of any knowledge about other agents' honesty and trustworthiness, the weights can be assumed to be 1.

3.4.3 Risk Attitude and Deceivability

A dual of the trustworthiness in the trust model, a *deceivability* is maintained by each sender to each receiver, which shows how easily the receiver can be cheated as seen by the sender. We define *deceivability*, c_{ij}, of receiver *j* from sender *i*'s point of view, as a function of the impression that agent *i* has about agent *j*, agent *i*'s calculation on the reputation of agent *j* as well as the risk attitude of agent *i*, which returns a real number in $[-1, 1]$:

$$c_{ij} = f_c(imp_{ij}, rep_{ij}, r_i)$$

The function f_c must satisfy a similar set of axioms for function f_t as stated in section 3.2.3:

Axiom f_{c1}: f_c is continuous.

Axiom f_{c2}: f_c decreases as imp_{ij} decreases and vice versa.

Axiom f_{c3}: f_c decreases as rep_{ij} decreases and vice versa.

Axiom f_{c4}: f_c decreases as r_i decreases and vice versa.

Axioms f_{c2} and f_{c3} state that it is rational that the deceivability of the receiver decreases if the sender's impression on it decreases, or the receiver's reputation decreases and vice versa. If the risk attitude of the sender decreases, which implies that the sender becomes more risk-averse and thus less willing to cheat other agents, then the evaluated deceivability of the receiver will decrease. This is axiom f_{c4}.

An example function satisfying the above axioms and the intuitive meanings is shown below, which attaches the same degree of importance to impression and reputation, and is in proportion to the agent's risk attitude.

$$c_{ij} = \frac{imp_{ij}+rep_{ij}}{2} \times (1 - r_i)$$

3.4.4 Temptation of Lying vs. Sincerity of the Sender

For a sender to decide whether to tell a lie, besides considering the expected payoffs that the agent can gain by lying, it should also consider the deceivability of the receiver. Formally, a sender makes use of a *temptation function* f_{tp} to calculate the temptation of lying. The *temptation*, t_L, of a lie L is defined by:

$$t_L = f_{tp}(r_i, c_{ij}, u_k)$$

Intuitively, the function f_{tp} takes the risk attitude r_i of sender i as the first argument, the deceivability c_{ij} of receiver j as seen by agent i as the second argument and the expected increase in utility u_k as the last argument, and returns a real number in $[-1, 1]$ as the temptation of lying. The function f_{tp} must satisfy a similar set of axioms for function f_p as stated in section 3.2.4:

Axiom f_{tp1}: f_{tp} is continuous.

Axiom f_{tp2}: (*Adventurousness of risk-seeking agents*) There exists a value $r_0 \in \Re$ such that $f_{tp}(r, c_2, u_1) > f_{tp}(r, c_1, u_2)$ if and only if $r > r_0$, $c_1 > c_2$ and $u_1 > u_2$.

Axiom f_{tp3}: (*Cautiousness of risk-averse agents*) There exists a value $r'_0 \in \Re$ such that $f_{tp}(r, c_2, u_1) < f_{tp}(r, c_1, u_2)$ if and only if $r < r'_0$, $c_1 > c_2$ and $u_1 > u_2$.

Axiom f_{tp4}: if $u_1 \geq u_2$, $f_{tp}(r, c, u_1) \geq f_{tp}(r, c, u_2)$ for $r > r_0$ and $f_{tp}(r, c, u_1) \leq f_{tp}(r, c, u_2)$ for $r < r'_0$.

Axiom f_{tp5}: if $c_1 \geq c_2$, $f_{tp}(r, c_1, u) \geq f_{tp}(r, c_2, u)$.

Axiom f_{tp6}: if $r_1 \geq r_2$, $f_{tp}(r_1, c, u) \geq f_{tp}(r_2, c, u)$.

Axiom f_{tp2} and f_{tp3} state that it is rational for a risk-seeking sender to consider expected gain in utility to be more important than deceivability of the receiver, and vice versa to a risk-averse sender. At the same time, temptation of lies that bring more utility should be higher for a risk-seeking sender, but lower for a risk-averse sender, as it is rational for a risk-averse sender to be hesitate to tell a lie with higher utility. This brings about axiom f_{tp4}. In addition, the temptation of lying a more deceivable receiver should be higher, which is axiom f_{tp5}. However, the temptation of lying decreases if the sender becomes more risk-averse, which is axiom f_{tp6}.

A simple example satisfying the above axioms, f_{tp} can be defined as follows:[3]

$$f_{tp}(r, c, u) = \begin{cases} \frac{(r-1)u+c}{2} & \text{for } r < 0.5 \\ \frac{u+c}{2} & \text{for } r = 0.5 \\ \frac{(r+1)u+c}{3} & \text{for } r < 0.5 \end{cases}$$

With this function, the more risk-averse the sender is, the more important the deceivability of the receiver is in making decision. The more risk-seeking the sender

[3]In this formula, the utility is assumed to be in the range $[0, 1]$.

is, the more important the utility of the lie is, while senders with neutral risk attitude consider deceivability of the receiver and utility of the lie to be the same important.

A sender should decide to tell a lie only if the temptation of lying is greater than a certain threshold. We call this the threshold the *sincerity* of the sender to the receiver, which is a real number in $[-1, 1]$. Each sender maintains a sincerity to each receiver, which can change over time. If more than one lie can be chosen from, the sender should send the lie with the greatest temptation, among those lies having a temptation higher than the corresponding sincerity to the receivers. Since agents can only choose one action in each round, and the aim of lying is to change the competitor's action so as to make its own action compatible, agents will only choose at most one lie to send.

The function f_{tp} also have a set of theorems similar to that stated in section 3.2.4:

Theorem 6. *r_0 (in axiom f_{tp2}) equals r'_0 (in axiom f_{tp3}).*

Proof. Assume r_0 is not equal to r'_0, this results in the following two cases:

Case 1. $r'_0 > r_0$ By axiom f_{tp2}, if $r > r_0$, $c_1 > c_2$ and $u_1 > u_2$, $f_{tp}(r, c_2, u_1) > f_{tp}(r, c_1, u_2)$. By axiom f_{tp3}, if $r < r'_0$, $c_1 > c_2$ and $u_1 > u_2$, $f_{tp}(r, c_2, u_1) < f_{tp}(r, c_1, u_2)$. As a result, for $r_0 < r < r'_0$, $f_{tp}(r, c_2, u_1) > f_{tp}(r, c_1, u_2)$ and $f_{tp}(r, c_2, u_1) < f_{tp}(r, c_1, u_2)$, which is a contradiction.

Case 2. $r'_0 < r_0$ By axiom f_{tp2}, $f_{tp}(r, c_2, u_1) > f_{tp}(r, c_1, u_2)$ if and only if $r > r_0$, $c_1 > c_2$ and $u_1 > u_2$, so for $r < r_0$, $f_{tp}(r, c_2, u_1) \leq f_{tp}(r, c_1, u_2)$. By axiom f_{tp3}, $f_{tp}(r, c_2, u_1) < f_{tp}(r, c_1, u_2)$ if and only if $r < r'_0$, $c_1 > c_2$ and $u_1 > u_2$, so for $r > r'_0$, $f_{tp}(r, c_2, u_1) \geq f_{tp}(r, c_1, u_2)$. As a result, for $r'_0 < r < r_0$, $f_{tp}(r, c_2, u_1) \leq f_{tp}(r, c_1, u_2)$ and $f_{tp}(r, c_2, u_1) \geq f_{tp}(r, c_1, u_2)$, which is a contradiction.

So r_0 equals r'_0. □

Theorem 7. *If two lies M_1 and M_2 are available to all senders with different risk attitudes, while M_1 and M_2 have expected utilities u_1, u_2 and deceivability of receivers c_1, c_2, respectively, where $u_1 > u_2$ and $c_2 > c_1$. Then there exists a constant*

$r_0 \in \Re$ depending only on u_1, u_2 and c_1, c_2, such that all the senders with risk attitude $r > r_0$ will choose to send out M_1 and all the senders with risk attitude $r < r_0$ will choose to send out M_2 if temptation of the lies are greater than the senders' sincerity.

Proof. Sender uses a function f_{tp} to rank the lies and choose to tell the lie that has the highest value of f_{tp}, where f_{tp} must satisfy axioms f_{tp1} to axioms f_{tp6}. Since $c_2 > c_1$ and $u_1 > u_2$, by axiom f_{tp2}, there exists a value $r_0 \in \Re$ such that $f_{tp}(r, c_1, u_1) > f_{tp}(r, c_2, u_2)$ if and only if $r > r_0$. And by axiom f_{tp3}, there exists a value $r'_0 \in \Re$ such that $f_{tp}(r, c_1, u_1) < f_{tp}(r, c_2, u_2)$ if and only if $r < r'_0$. By Theorem 6, $r = r'_0$. So, for $u_1 > u_2$ and $c_2 > c_1$, there exists a value $r_0 \in \Re$ such that if $r > r_0$, then $f_{tp}(r, c_1, u_1) > f_{tp}(r, c_2, u_2)$, which means the sender will choose to tell lie M_1, and if $r < r_0$, then $f_{tp}(r, c_1, u_1) < f_{tp}(r, c_2, u_2)$, which means the sender will choose to tell lie M_2 if temptation of the lies are greater than the senders' sincerity. □

Theorem 8. *Suppose there are two senders R_1 and R_2, with risk attitudes r_1 and r_2 respectively, where $r_1 > r_2$, that is sender R_1 is more risk-seeking than sender R_2. Then there exist two lies M_1 and M_2, with expected utilities u_1, u_2 and deceivability of receivers c_1, c_2, respectively, where $u_1 > u_2$ and $c_2 > c_1$, such that when these two messages are available to R_1 and R_2, R_1 will choose to send message M_1 and R_2 will choose to send message M_2 if temptation of the lies are greater than the senders' sincerity.*

Proof. By theorem 7, for any two lies M_1 and M_2, with expected utilities u_1, u_2 and deceivability of receivers c_1, c_2, respectively, where $u_1 > u_2$ and $c_2 > c_1$, there exists a constant $r_0 \in \Re$ depending only on u_1, u_2 and c_1, c_2, such that all the senders with risk attitude $r > r_0$ will choose to tell lie M_1 and all the senders with risk attitude $r < r_0$ will choose to tell lie M_2 if temptation of the lies are greater than the senders' sincerity. In other words, proving theorem 8 is to find the two lies such that $r_2 < r_0 < r_1$. We do this by first initialize two lies M_1 and M_2, with expected utilities u_1, u_2 and deceivability of receivers c_1, c_2, respectively, where $u_1 > u_2$ and

$c_2 > c_1$. When these two lies are available to the senders, one of the following four cases will result:

Case 1. Both of R_1 and R_2 choose to tell lie M_1. In this case, generate another two messages M_1' and M_2', with expected utilities u_1', u_2' and deceivability of receivers c_1', c_2', respectively, where $u_1' > u_2'$, $c_2' > c_1'$, $c_2' > c_2$ and $c_1' < c_1$.

Case 2. Both of R_1 and R_2 choose to tell lie M_2. In this case, generate another two messages M_1' and M_2', with expected utilities u_1', u_2' and deceivability of receivers c_1', c_2', respectively, where $u_1' > u_2'$, $c_2' > c_1'$, $u_1' > u_1$ and $u_2' < u_2$.

In case 1 and case 2, the process is continued by replacing the old lies M_1 and M_2. As r is a real number, as long as $r_1 > r_2$, there exists $r_0 \in \Re$, such that $r_2 < r_0 < r_1$. So, eventually, the process converge and case 3 will results.

Case 3. R_1 chooses to tell lie M_1 and R_2 chooses to tell lie M_2. In this case, the theorem is proved.

Case 4. R_1 chooses to tell lie M_2 and R_2 chooses to tell lie M_1. In fact, this case will never happen. Suppose R_1 and R_2 choose to tell different lies, as $r_1 > r_2$, and by theorem 2, $r_2 < r_0 < r_1$, which means R_1 will choose to tell lie M_1 and R_2 will choose to tell lie M_2. □

It is rational for a risk-seeking sender to send out a lie with a higher utility to a more deceivable receiver, rather than a lie with a lower utility and to a less deceivable receiver. For a risk-averse sender, as it will be hesitate to tell a lie with a higher utility, which lie it chooses to send depends on the actual values of the deceivability, utilities, risk attitude and sincerity. This is represented by the following theorem.

Theorem 9. *For risk attitude $r > r_0$, deceivability c_1 and c_2, and utilities u_1 and u_2, where $c_1 \geq c_2$ and $u_1 \geq u_2$, $f_{tp}(r, c_1, u_1) \geq f_{tp}(r, c_2, u_2)$.*

Proof. From axiom f_{tp4}, if $u_1 \geq u_2$, $f_{tp}(r, c_1, u_1) \geq f_{tp}(r, c_1, u_2)$ for $r > r_0$. From axiom f_{tp5}, if $c_1 \geq c_2$, $f_{tp}(r, c_1, u_2) \geq f_{tp}(r, c_2, u_2)$. So, $f_{tp}(r, c_1, u_1) \geq f_{tp}(r, c_1, u_2) \geq f_{tp}(r, c_2, u_2)$. That is $f_{tp}(r, c_1, u_1) \geq f_{tp}(r, c_2, u_2)$. □

The following theorem confirms that if a sender becomes more risk-averse and lowers the deceivability of the receiver after it fails to cheat the receiver, then it should be less willing for the sender to tell the same lie to the same receiver (with deceivability lowered).

Theorem 10. *For risk attitudes r_1 and r_2, deceivability c_1 and c_2, and utility u, where $r_1 \geq r_2$ and $c_1 \geq c_2$, $f_{tp}(r_1, c_1, u) \geq f_{tp}(r_2, c_2, u)$.*

Proof. From axiom f_{tp6}, if $r_1 \geq r_2$, $f_{tp}(r_1, c_1, u) \geq f_{tp}(r_2, c_1, u)$. From axiom f_{tp5}, if $c_1 \geq c_2$, $f_{tp}(r_2, c_1, u) \geq f_{tp}(r_2, c_2, u)$. So, $f_{tp}(r_1, c_1, u) \geq f_{tp}(r_2, c_1, u) \geq f_{tp}(r_2, c_2, u)$. That is $f_{tp}(r_1, c_1, u) \geq f_{tp}(r_2, c_2, u)$. □

3.5 Duality of the Trust/Honesty Model

The Trust Model enables receivers to decide whether or not to believe the received message(s), while the Honesty Model enables senders to decide whether or not to lie. In fact, the Honesty Model for the senders is a dual of the Trust Model for the receivers. In both models, receivers and senders maintain impression and reputation of senders and receivers respectively. In the Trust Model, receivers maintain trustworthiness of the senders, while senders maintain deceivability of the receivers in the Honesty Model. For a receiver to determine whether to believe the received message(s) with the Trust Model, persuasiveness of the messages are compared with stubbornness of the receiver. For a sender to determine whether to lie with the Honesty Model, temptation of lying is compared with sincerity of the sender. As deceivability is a dual of trustworthiness, and persuasiveness is a dual of temptation, the functions share similar sets of axioms and theorems.

3.6 Performance of the Trust/Honesty Model

Simulations are done to compare performance of agents employing our Trust/Honesty Model with performance of agents adopting other models or strategies. The setting of the simulation is as follows. We include receivers and senders adopting our Trust/Honesty Model. In addition, we include receivers and senders adopting other models or strategies. For receivers adopting our Trust/Honesty Model, a negative stubbornness and a risk attitude of 0.2 are used. For receivers adopting Sabater and Sierra's REGRET Model [SS01] and Mui *et al.*'s Computational Model of trust and reputation [MMH02], reputation and trust are calculated with the parameters suggested in these papers. These two receivers choose to believe the message from a sender with the maximum reputation when several messages are received at a time. If only one message is received, receiver adopting Mui *et al.*'s Computational Model chooses to believe the message if the sender's reputation is greater than 0.5, as suggested in the paper [MHM02]. Since Sabater and Sierra have not suggest any threshold and 0.5 is general enough to be a threshold, receiver adopting Sabater and Sierra's REGRET Model also chooses to believe the message if the sender's reputation is greater than 0.5 when only one message is received. Receivers adopting the "Choose Maximum Reputation" strategy chooses to believe the message from a sender with the maximum reputation when several messages are received at a time, where the reputation is calculated as suggested in this chapter. Using this calculation, an agent may have negative reputation. So, if only one message is received, this receiver chooses to believe the message if and only if the reputation of the message sender is positive. Similarly, receiver adopting the "Choose Maximum Utility" strategy choose the message with maximum utility to believe when several messages are received at a time, and chooses to believe the message if the utility of the message is greater than 0.5, where the utility is normalized to 1. Finally, receivers adopting the Random strategy randomly choose to believe a message when several messages are received at a time, and randomly choose to believe or not to believe the message when only one message is received.

In each round, a random semi-competitive scenario is virtually generated. Each sender decides whether to tell a lie to a receiver according to its adopted strategy. Therefore, it is possible that a receiver receives more than one message at a time.

Models/strategies	Utility gain
Maximum possible utility	1980
Trust/Honesty Model	1502
Sabater and Sierra's REGRET Model [SS01]	521
Choose Maximum reputation	508
Mui *et al.*'s Computational Model [MMH02]	499
Random	-724
Choose Maximum Utility	-812

Table 3.4: Average utility gain of receivers

Each receiver then chooses whether to believe the message according to its adopted strategy. Note that a receiver adopting the Trust/Honesty Model may believe no message at all if the persuasiveness of the messages it receives are all less than its stubbornness. At the end of each round, a receiver gains if it has believed a true message, or loses if it has believed a lie. On the other hand, a sender gains if the receiver has believed its message, or loses if the receiver has not. Then all agents update the impressions, reputations, trustworthiness, and deceivability accordingly. In these simulations, all agents' risk attitudes, stubbornness values, and sincerity values do not change throughout the game. Each game contains 5,000 rounds, and the average results of 1,000 games are shown in Table 3.4 and Table 3.5.

Table 3.4 shows the average utility gain[4] of receivers. In the table, maximum possible utility means the maximum utility a receiver can possibly gain if it is so smart as to always choose the right message to believe, and has never been cheated. Note that this just serves as a benchmark for the comparison. Experiments show that receivers adopting our Trust/Honesty Model significantly outperform the others by at least 3 times. This is because the REGRET Model and Mui *et al.*'s Computational Model do not take utility into account in making decisions. Utility of the receiver adopting the "Choose Maximum Reputation" strategy is similar to those of

[4]Rounded up to the nearest integer.

Models/strategies	Utility gain
Maximum possible utility	1868
Trust/Honesty Model	1230
100% Truth	616
Mui *et al.*'s Computational Model [MMH02]	540
Sabater and Sierra's REGRET Model [SS01]	501
Random 50% Truth	-678
Always Lie	-1769

Table 3.5: Average utility gain of senders

the receivers adopting Sabater and Sierra's REGRET Model and Mui *et al.*'s Computational Model. The receivers adopting the Choose Maximum Utility strategy and the Random strategy end up with negative utilities, because they are easily cheated.

Table 3.5 shows the average utility gain of senders. Again in the table, maximum possible utility means the maximum utility that a sender can possibly get if it can always gain receivers' trust. Again, this serves only as a benchmark for comparison. Among all senders, sender adopting our Trust/Honesty Model with risk attitude 0.4 and sincerity 0.8 has the highest utility. The sender adopting the 100% Truth strategy always tells the truth. However, receivers may not believe it if the utility brought by the messages is not attractive, so its performance is not the best. Senders adopting the REGRET Model and Mui *et al.*'s model choose to tell lies if the target receiver has good reputation of being deceivable,. Their results are similar but not very good as they only take reputation into account, but do not consider utility in making decision. The sender adopting the Random 50% Truth strategy randomly tells 50% of truth and the one adopting the Always Lie strategy always tells lies. As a result, their utilities are negative, as their reputations are low and no receiver believe them.

These experiments show that our Trust/Honesty Model significantly outperforms other trust models. It helps agents to achieve a utility that is about two to three times better than that achieved by agents adopting other trust models reported in the literature. Experiments also show that considering only reputation or only expected utility cannot achieve high utility.

3.7 Summary

In semi-competitive environments, agents have intentions to be honest and have intentions to lie. This chapter introduces a Trust/Honesty Model for agents to choose another type of actions, which is whether to believe a received message, and whether to be honest. Simulations shown that agents with our Trust/Honesty Model significantly outperform agents with other existing models or strategies.

Chapter 4

Adaptive Strategies

In this chapter, we improve the Trust/Honesty Model to an adaptive one. We first introduce the problem of non-adaptive agents. Then we design the adaptive strategy. Finally, we compare the performance of adaptive agents with the non-adaptive ones.

4.1 Problem of Non-adaptive Agents

In receivers' trust model, there are two parameters that can be varied, which are risk attitude and stubbornness, like the personality of human beings. As symmetry, risk attitude and sincerity can also be varied in senders' honesty model. Non-adaptive agents means their risk attitude, stubbornness, and sincerity do not change throughout the game.

Simulation is done to compare performance of non-adaptive agents with different parameters. In the simulation, there are 66 receivers, with risk attitude 0, 0.2, ..., 1 and stubbornness -1, -0.8, ..., 1, interact with 66 senders, with risk attitude 0, 0.2, ..., 1 and sincerity -1, -0.8, ..., 1. In each round, a random semi-competitive scenario is virtually generated. Each sender calculates the temptation of lying and decides whether to tell a lie to a receiver. Each receiver then calculates the persuasiveness of each message and chooses to believe the message with the highest persuasiveness among those persuasiveness higher than its stubbornness. A receiver may believe no message if the persuasiveness of the messages are all lower

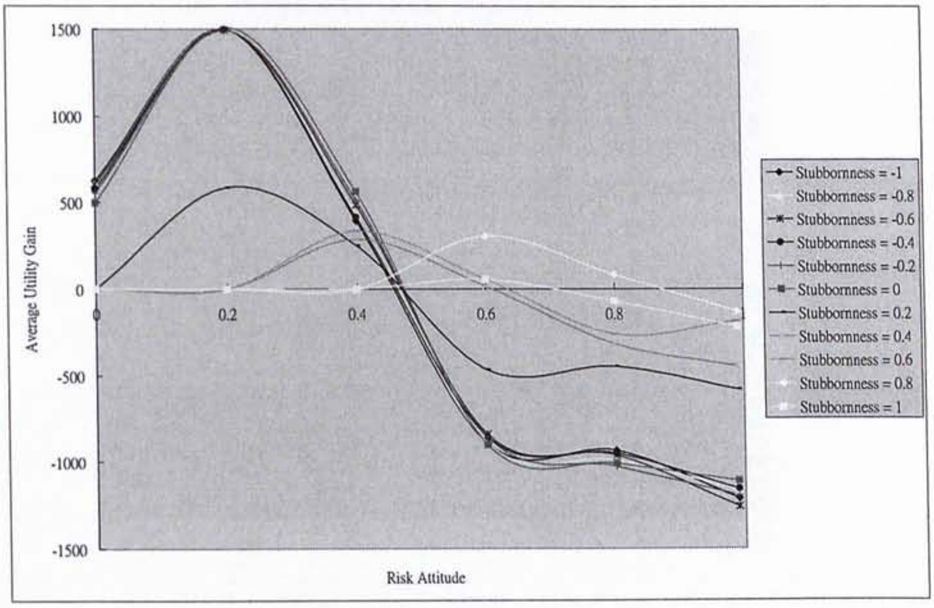

Figure 4.1: Average utility gain of non-adaptive receivers when interact with non-adaptive senders

than its stubbornness. At the end of each round, a receiver gains if it believed a true message, or loses if it has believed a lie. On the other hand, a sender gains if the receiver has believed its message, or loses if the receiver has not. Then all agents update the impressions, reputations, trustworthiness, and deceivability accordingly. In this simulation, all agents' risk attitude, stubbornness, and sincerity do not change throughout the game. Each game contains 5000 rounds, and the average results of 100 games are shown in Fig. 4.1 and Fig. 4.2.

Fig. 4.1 shows the average utility gain of non-adaptive receivers when interact with non-adaptive senders. From the figure, we see that receivers with negative stubbornness and risk attitude 0.2 perform the best. However, performance decreases for more risk-seeking receivers. On the other hand, for receivers with positive stubbornness, they cannot gain anything if they are risk-averse.

Fig. 4.2 shows the average utility gain of non-adaptive senders when interact with non-adaptive receivers. From the figure, less sincere senders have negative utility gain. This is because these senders always lie, so receivers' impressions on them and their trustworthiness are very low, which means receivers seldom believe

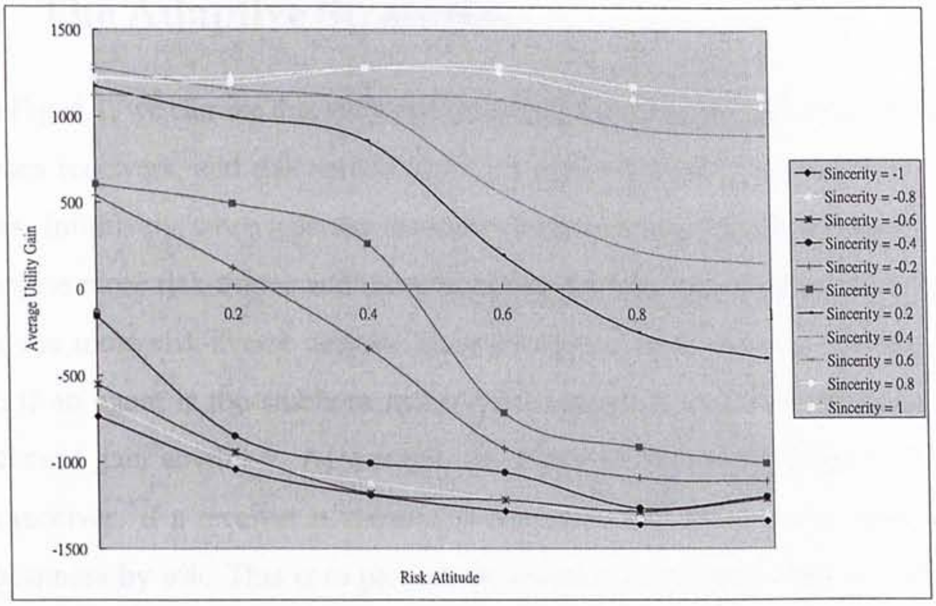

Figure 4.2: Average utility gain of non-adaptive senders when interact with non-adaptive receivers

them. On the other hand, sincere senders have the best performance. Also, performance decreases for more risk-seeking senders.

From Fig. 4.1 and Fig. 4.2, we can see that the performance of non-adaptive agents depends very much on their risk attitudes, stubbornness and sincerity. In addition, choices of parameters depend on the type of interacting agents. For example, if a sender knows the receiver, that it is interacting with, is very stubborn, then the sender needs to be more sincere. In contrast, if a sender knows the receiver, that it is interacting with, is very risk-seeking and not stubborn at all, which means the receiver can be cheated very easily, then the sender can be less sincere and more risk-seeking. However, in real practice, receivers(senders) can hardly find out the type of senders(receivers) that they are interacting with. So, agents need to be adaptive.

4.2 The Adaptive Strategies

From Fig. 4.1, we can see that stubborn receivers in general perform better than less stubborn receivers, and risk neutral receivers perform better than risk-seeking receivers. Intuitively, when a person discovers that he is cheated, it is rational for him to become more risk-averse and more stubborn. In addition, the more the receiver loses, the more risk-averse and the more stubborn it will become. On the other hand, if an agent is too stubborn and too risk-averse, it may believe no message and cannot gain anything. As a result, we derive the following adaptive strategy for a receiver: if a receiver is cheated, it lowers its risk attitude and increases its stubbornness by $p\%$. This is to prevent the receiver from being cheated again. At the same time, if it has not believed any message for more than m rounds, than it increases its risk attitude and lowers its stubbornness by $\frac{1}{p}\%$. This is to prevent a receiver from being too risk-averse and too stubborn so as to prevent a receiver from believing no message and gain nothing. Note that m cannot be too large, otherwise the function will be lose. Also, m cannot be too small. This is because a receiver can choose not to believe the received message if the message is not persuasive enough, in this way, it needs not change its risk attitude or stubbornness. In addition, it is rational that an agent becomes more risk-seeking and less stubborn if it has believed the true messages for some rounds. So if a receiver has believed the true messages for more than k rounds, it increases its risk attitude and lowers its stubbornness by $\frac{1}{p}\%$. However, k cannot be too small, otherwise, the agent will become too risk-seeking. In the following simulations, each game consists of 5000 rounds, we use $m = k = 10$. The adaptive percentage $p = \frac{accloss_n}{accgain_n}$. $accloss_n = \alpha \times accloss_{n-1} + (1-\alpha) \times loss_n$, which is a weighted sum of the utility loss in the past and the utility loss in the latest round. It is rational that the latest loss will have the most influence and the oldest loss will have the least influence, so $0 \leq \alpha \leq \frac{1}{2}$. Similarly, $accgain_n = \alpha \times accgain_{n-1} + (1-\alpha) \times gain_n$. $loss_n$ is the amount that the agent loses in the latest round and $gain_n$ is the amount that the

combination	risk attitude	stubbornness/sincerity
1	kept constant	kept constant
2	kept constant	changes over time
3	changes over time	kept constant
4	changes over time	changes over time

Table 4.1: Combinations on variations of parameters of the Trust/Honesty Model

agent gains in the latest round. In particular, $p = 1$ if $accgain_n = accloss_n = 0$, $p = 100$ if $accgain_n = 0$ but $accloss_n \neq 0$, and $p = \frac{1}{100}$ if $accloss_n = 0$ but $accgain_n \neq 0$.

From Fig. 4.2, we can see that less sincere senders have negative utility gain and sincere senders perform much better. For sincere senders, performance depends less on the risk attitude. However, for less sincere senders, performance decrease as risk attitude increase. Intuitively, when a person discovers that he is less trusted by others, it is rational for him to become more risk-averse and more sincere. As a result, we derive the following adaptive strategy for a sender: if no receiver believes the sender in the latest round, the sender lowers its risk attitude and increases its sincerity by $p\%$. This is to prevent the sender from telling too many lies. At the same time, if it does not send out any message for more than m rounds or it has gained the receiver's trust for more than k rounds, it increases its risk attitude and lowers its sincerity by $\frac{1}{p}\%$. This is to prevent the sender from being too risk-averse to send out any message and gain nothing.

4.3 Variations of Parameters

In the Trust/Honesty Model, there are three parameters which can be varied: risk attitude, stubbornness and sincerity. In fact, there are four combinations that the parameters can be varied, which is shown in Table 4.1. Let us look at the intuitive meaning of the four combinations of variations. Consider a sender, which sends the same lie to the same receiver in every round. Every time the receiver is cheated, it lowers the trustworthiness of the sender, adjusts its own risk attitude and stubbornness accordingly.

For the first combination, receiver is non-adaptive, and its risk attitude and stubbornness are kept constant even it is cheated. Every time the receiver keeps receiving the same lie, as utility brought by the message is the same and its risk attitude is constant, persuasiveness of the message will be lowered as trustworthiness of the sender is lowered. Eventually, persuasiveness of this message will fall below the stubbornness, and the receiver will stop being cheated. In this way, the rate that the receiver learns from its experience is proportional to the rate that the trustworthiness of the sender is decreased. For the second combination, where risk attitude is kept constant and stubbornness is increased as a reaction to the lie, the rate that the receiver learns its experience is proportional to the rate that the trustworthiness of the sender is decreased plus the rate that the stubbornness is increased. For the third combination, where risk attitude changes over time and stubbornness is kept constant, the rate that the receiver learns its experience is proportional to the rate that the trustworthiness of the sender is decreased plus the rate that the risk attitude is decreased.

The last combination, where both risk attitude and stubbornness changes over time, is the adaptive strategy. After the first time the receiver is cheated, it lowers the trustworthiness of the sender. The second time the receiver receives the same lie, as utility of the message is the same and both its risk attitude and trustworthiness of the sender are lowered, by Theorem 5, persuasiveness of the message will be lowered. At the same time, stubbornness is increased as a reaction to the lie. Eventually, this persuasiveness will fall below the stubbornness, and the receiver will stop being cheated. In this way, the rate that the receiver learns its experience is proportional to the rate that the trustworthiness of the sender is decreased plus the rate that the risk attitude is decreased plus the rate that the stubbornness is increased. Note that if the receiver, which has become more risk-averse, receives a message

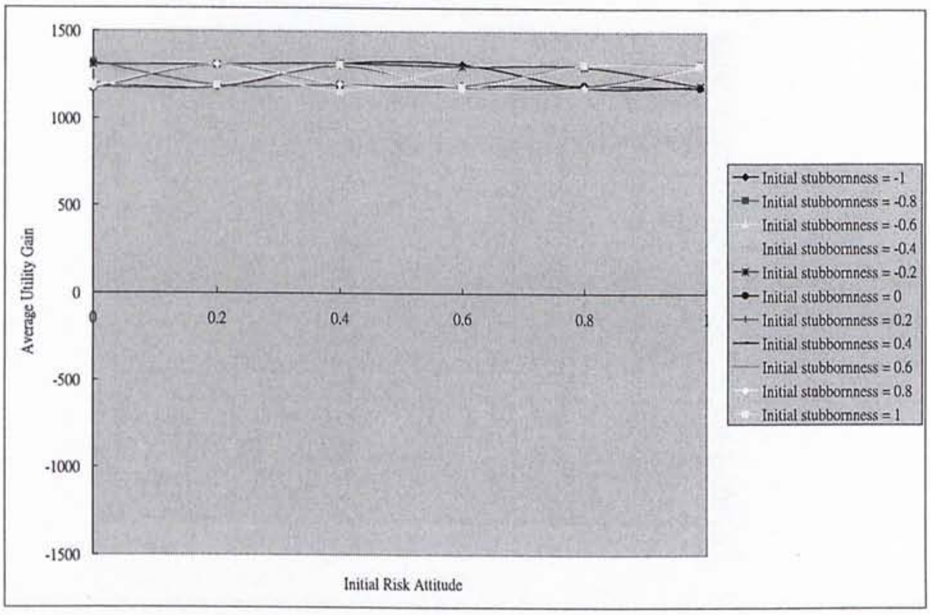

Figure 4.3: Average utility gain of adaptive receivers when interact with non-adaptive senders

with higher payoff, from the same sender with lower trustworthiness, we cannot determine whether the persuasiveness of this message is higher than the previous one, as that depends on the how high the payoff is.

To conclude, keeping both risk attitude and stubbornness unchanged (non-adaptive) have the slowest rate of learning and varying both (adaptive) have the fastest rate of learning the experience.

4.4 Adaptive Agents vs. Non-adaptive Agents

Simulations are done to compare the performance of adaptive agents with that of non-adaptive agents. In these simulations, both the adaptive agents and the non-adaptive agents adopt the Trust/Honesty Model. Simulation settings are the same as section 4.1. In Chapter 6, we will compare the performance agents adopting different models and strategies.

Fig. 4.3 shows the average utility gain of adaptive receivers when interact with non-adaptive senders. From the figure, we can see that all receivers can attain a

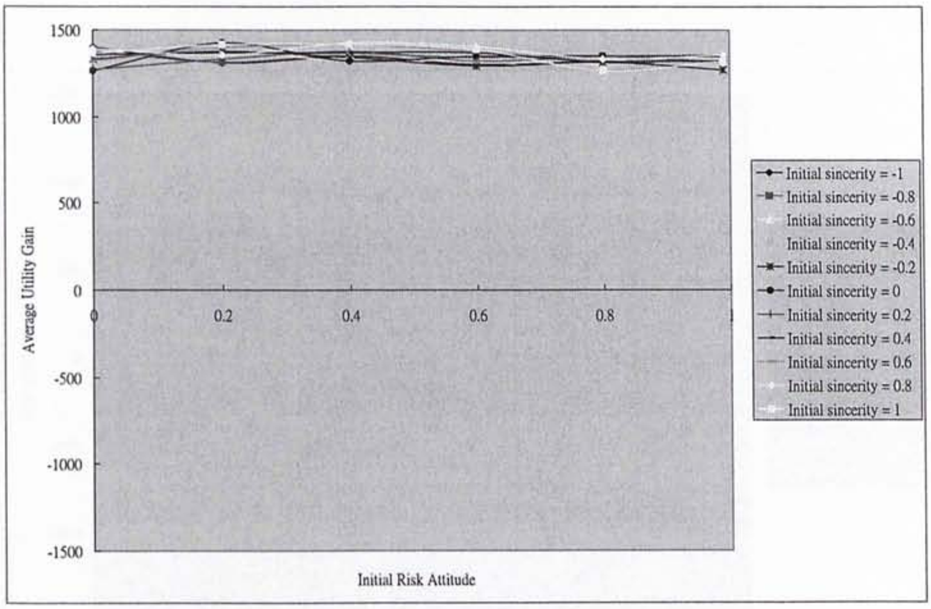

Figure 4.4: Average utility gain of adaptive senders when interact with non-adaptive receivers

utility, which is very close to the maximum utility in Fig. 4.1. Also, utility gain becomes independent of the initial choices of risk attitude and stubbornness. Fig. 4.4 shows the average utility gain of adaptive senders when interact with non-adaptive receivers. From the figure, we can see that all senders can attain a utility, which is very close to the maximum utility in Fig. 4.2. Also, utility gain becomes independent of the initial choices of risk attitude and sincerity. In this way, the problem described in section 4.1 is solved.

Fig. 4.5 shows the average utility gain of non-adaptive receivers when interact with adaptive senders. Compare the figure with Fig. 4.1, we can see that when senders become adaptive, which means senders learn to tell less lies when they discover that they are less trusted by the receivers, receivers' utility gain increase in general. Only the utility gain of less stubborn receivers decrease a little bit, that is because senders adapt to these receivers and tell more lies to them. Fig. 4.6 shows the average utility gain of non-adaptive senders when interact with adaptive receivers. Compare the figure with Fig. 4.2, we can see that when receivers become adaptive, which means receivers learn to be more risk-averse and more stubborn

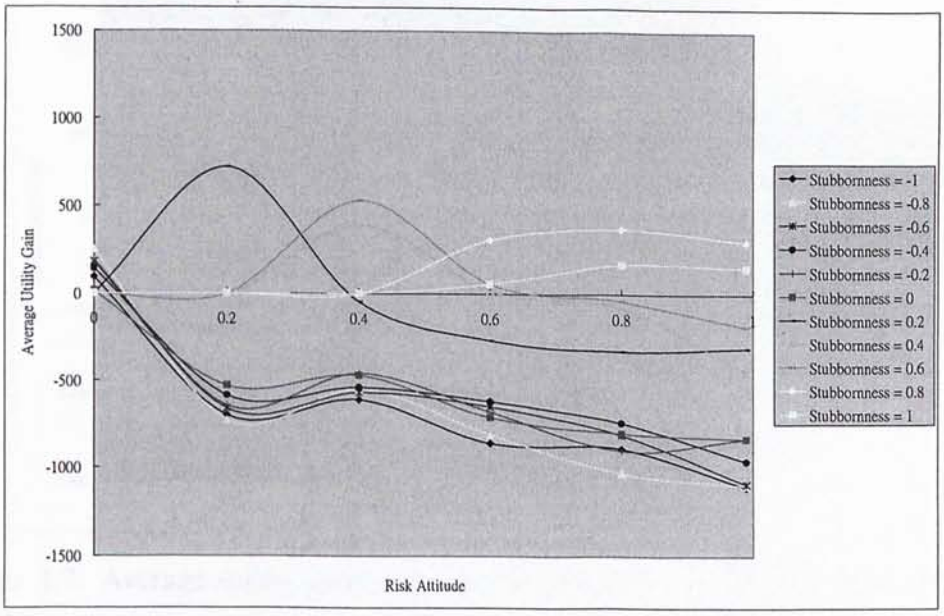

Figure 4.5: Average utility gain of non-adaptive receivers when interact with adaptive senders

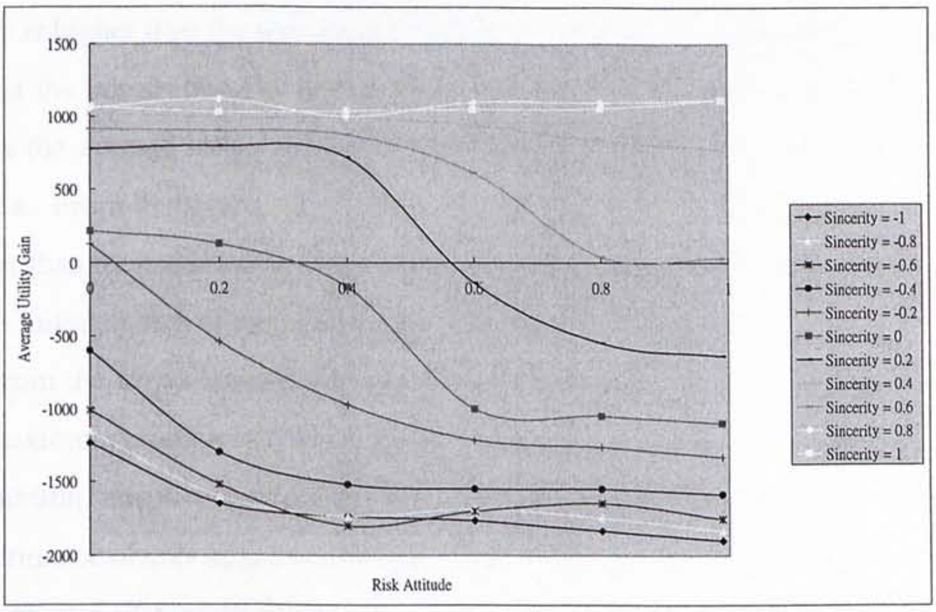

Figure 4.6: Average utility gain of non-adaptive senders when interact with adaptive receivers

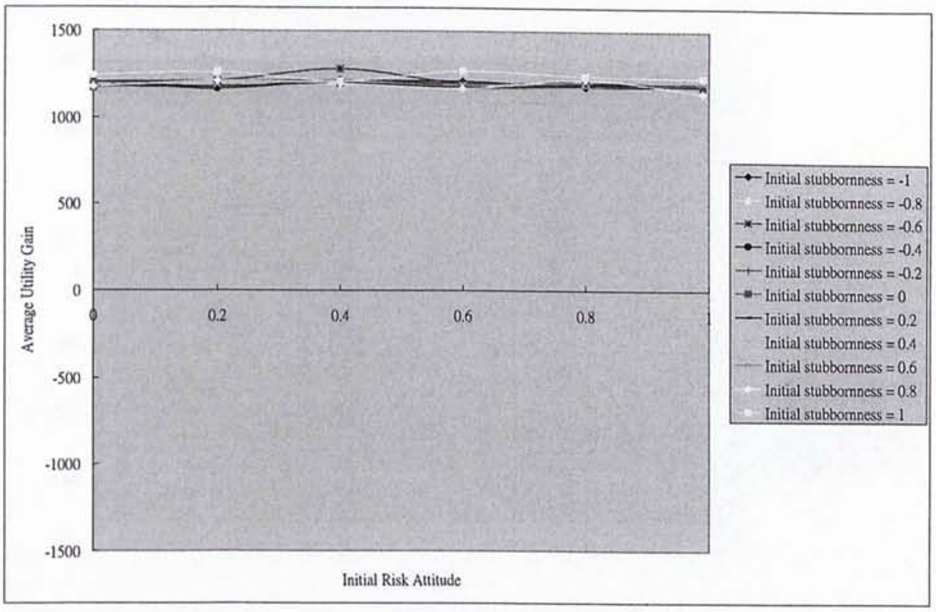

Figure 4.7: Average utility gain of adaptive receivers when interact with adaptive senders

when they are cheated, so senders' utility gain decrease in general.

Fig. 4.7 shows the average utility gain of adaptive receivers when interact with adaptive senders. From the figure, we can see that all receivers can attain a utility, which is higher than the maximum utility in Fig. 4.7 and the utility gain is independent of the initial choice of risk attitude and stubbornness of the receivers. Fig. 4.8 shows the average utility gain of adaptive senders when interact with adaptive receivers. From the figure, we can see that all senders can attain a utility, which is higher than the maximum utility in Fig. 4.8 and the utility gain is also independent of the initial choice of risk attitude and sincerity of the senders.

From the above simulation results, adaptive agents can attain a utility close to the maximum possible utility, no matter the opponents are adaptive or non-adaptive. In addition, adaptive agents do not need initial choices on parameters. However, performance of non-adaptive agents depends on the type of interacting agents. Also, performance of non-adaptive agents depends very much on the initial choices of parameters. So, we can conclude that adaptive agents perform better than the non-adaptive agents.

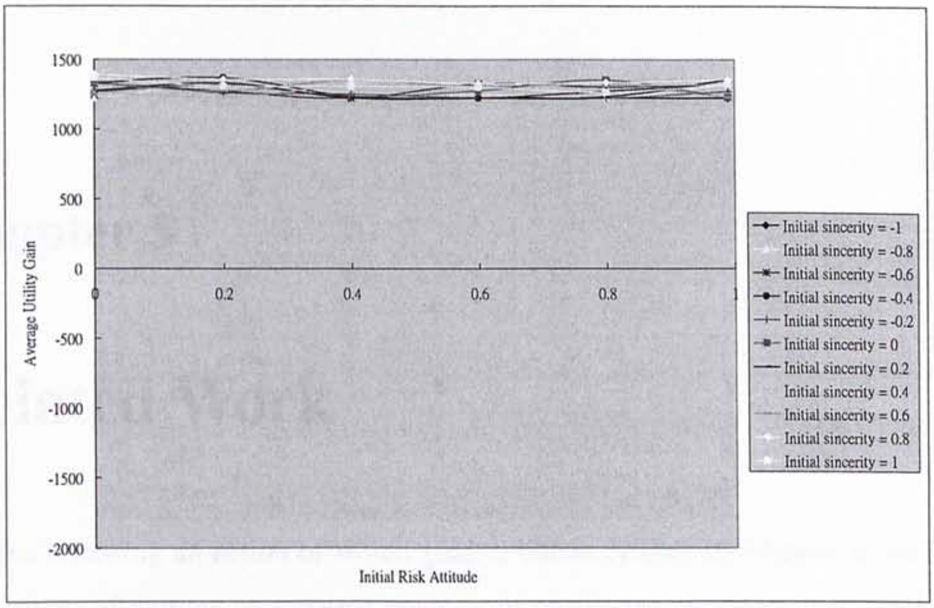

Figure 4.8: Average utility gain of adaptive senders when interact with adaptive receivers

4.5 Summary

This chapter presents an adaptive strategy in addition to the Trust/Honesty Model, which allows agents to adapt to the environment and improves the problem of the non-adaptive models. Simulations show that adaptive agents perform better than the non-adaptive ones.

Chapter 5

Related Work

Besides choosing an action of which goal to obtain or choose whether to stay still, other types of actions that agents many need to choose are whether to trust other agents, and whether to be honest. In choosing these types of actions, agents need additional information, like impression, reputation, and trust. In the literature, there have been various definitions and representations of impression, reputation, and trust. In this chapter, we give a review of this topic. In addition, we give a review on the theory of honesty.

5.1 Impression, Reputation and Trust

Marsh [Mar94] is among the first researchers to introduce a computational model for trust. He defines *General Trust*, the amount of trust that agent x has in agent y, which is independent of the situation, as a real number between -1 and $+1$, where -1 means complete distrust and $+1$ means absolute trust. Marsh uses an estimation of the general trust, an agent-subjective measure about the importance of the situation, as well as utility to estimate the *Situational Trust*, which is the amount of trust that agent x has in agent y in a particular situation. However, he does not mention anything about reputation. Marsh mentions that an agent decides to cooperate with a particular agent in a particular matter, if the trust it has on that particular agent in that particular matter is greater than a cooperation threshold, which is calculated from the risk and importance of the matter, as well as the competence of that particular agent on that particular matter. The problem is that risk of a matter and competence of an agent on a particular matter are difficult to estimate in real practice. In addition, the framework is incomplete, as the way in which trust can be modified is not defined.

Mui *et al.* [MMH02, MMA$^+$01] use a Bayesian approach in the computational model of trust and reputation, in which they estimate the reputation of agent x in the eye's of agent y as the probability that agent x cooperates with agent y, that is the number of cooperation that agent x has made toward agent y out of the previous encounters. The reputation defined there is an opinion that a single agent has about a particular agent, rather than the opinion that a group of agents have about a particular agent. This deviate from the definitions in the dictionaries [htta, httb]. In the computational model, they define trust as the expected probability that agent x will cooperate the next time, given a history of encounters. There is a problem with this approach. Agents adopting this model can be cheated easily. For example, out of 10 encounters, agent x cooperates with agent y in 9 rounds bringing a utility gain of 10, but it does not cooperate in 1 round bringing a utility loss of 100, the expected probability that agent x will cooperate the next time is 0.9, so agent y will still trust agent x. However, agent x actually brings much more harm than gain to agent y, that is agent y is trusting a harmful agent. The reason for this is that this model only calculates the expected probability for cooperation, but does not include the utility that the interacting agent brings.

Mui *et al.* [MHM02] use the reputation model, called *Reputation Tic-for-tat*, to simulate the Prisoner's Dilemma game [Axe84]. Agent adopting traditional Tic-for-tat strategy, will cooperate initially, and then does what the other agent did (cooperate or defect) in the previous round. Agent adopting the Reputation Tic-for-tat strategy, will cooperate initially depending on the reputation of the other agent, and then does whatever the other agent did in the previous round. However, the

Reputation Tic-for-tat strategy is not much different from the traditional Tic-for-tat strategy. Note that this Reputation Tic-for-tat will be the same as the traditional tic-for-tat strategy if the reputation of agents is assumed to be high at the very beginning of the game when there is insufficient information. They define different types of reputation for the Reputation Tic-for-tat strategy. Among which, *propagated reputation* perform the best. Propagated reputation means when an agent encounter an unknown agent, it will ask other agents for the reputation of the unknown agent. In the experiment, there are only two types of agents. One of which always defect, named AllD agents. Another type use the Reputation Tic-for-tat strategy. When a Reputation Tic-for-tat agent encounter an unknown agent, what it need to do is just to identify whether the unknown agent is an AllD agent or a Reputation Tic-for-tat agent. Then defects the AllD agent and cooperates with the Reputation Tic-for-tat agent. In this way, it will always defect the AllD agents and always cooperate with the Reputation Tic-for-tat agents, and a maximum utility can be obtained. In fact, it is easy to make the identification because there are only two types of agents in the environment. Also, the AllD agents must have a very low reputation as they always defect and the Reputation Tic-for-tat agents must have a high reputation as they always cooperate with agents of the same type.

Sabater and Sierra [SS01] propose another reputation model. There they define *impression* that an agent has on another agent as the subjective evaluations made by an agent on certain aspects of the agent being evaluated, and they calculate *individual-experienced reputation* that an agent has on another agent, directly from an agent's impression database. For example, to evaluate the reputation of being a trustworthy sender, agent will consider the reputation of telling the truth. In this model, there is a *group-experienced reputation* that a group of agents have on a particular agent being evaluated. This is calculated by the weighted sum of the individual-experienced reputation that the member agents in the group has, on the agent being evaluated. This matches the definition from the dictionary [htta]. However, this work mainly concentrates on the calculation of impression and reputation, rather than showing how to use these information to make decisions.

Rubiera *et al.* [RLM01] also define reputation as the past experience of individual agent together with references from other agents. In addition, an agent will only choose some of the agents to ask references for and it will determine how much the received reference will count for. This is similar to a weighted sum of individual experience.

Yu and Singh [YS01, YS03] define reputation based on a probabilistic approach. For agent x to evaluate the trustworthiness of agent y, they calculate the reputation of agent y, which is done by combining the reputation of y as seen by a group of witness agents, as well as the reputation of the witness agents are integrated. These papers also concentrate on the calculation of impression and reputation as well as a network of trust information.

Castelfranchi and Falcone [CF98, FC01] describe the importance of trust and explain what trust is, though in a rather qualitative way. They also define under what situation should an agent delegate to other agents. In addition, they propose that risk should be taken into account when deciding whether to delegate, and agent should have a risk policy, which means agent should refuse a choice of decision if the hazard of that choice is greater than a certain threshold. This similar to our risk attitude. In addition, they have implemented their model to analyze the different nature of the belief sources and their trustworthiness [CFP03].

Glass and Grosz [GG00] use *Brownie Points* to represent an agent's historical reputation. The value of brownie points of an agent will be increased if the agent makes a socially conscious decision, and the value will be decreased otherwise. This representation measures the opinion that a group of agents in general have about a particular agent. This can help agents prevent lying, but in a passive way. In our proposed model, agents choose whether to lie and it can prevent itself from lying by the adaptive strategy.

Furthermore, Griffiths and Luck [GL03] apply the concept of trust in coalition formation, in which agents can benefit mutually. They define trust as a representation of an agent's estimation of how likely another agent is to fulfill its cooperative environment, which is inferred based on agents' experience over time. They also mention that the trust values can be updated according to the agents' personality: optimistic or pessimistic. However, they do not have a quantitative definition.

5.2 Theory of Honesty

The issues of honesty have also been addressed by Gmytrasiewicz and Durfee [GD93]. For an agent to decide whether to lie, they first model the respective actions that the receiver will take on believing and not believing the lie. Then they calculate the expected utilities on telling and not telling lie. To calculate the expected utility on telling lie, they consider the resulting utility if the receiver believes the lie and the resulting utility if the receiver does not believe the lie. An agent decides to lie only if the expected utility of lying is greater than that of being honest. In their model, only expected utility is considered in deciding whether to believe a message and deciding whether to lie. The problem is obvious, agents are easily cheated by those lies, which claim to bring high expected utilities.

5.3 Summary

In this chapter, we review previous research done on the calculation of trust, impression, reputation and honesty. Definitions on reputation [MHM02, SS01, RLM01, YS01, YS03] are similar. In addition to calculation and definition, these research also apply the trust information to decide whether to cooperate with other agents. Another type of decision-making is to decide whether to lie or whether to believe a message. In choosing such an action, some of the previous work consider only trust and/or reputation, while some only consider expected utility, but little consider both.

Chapter 6

Performance Analysis

In this chapter, simulation is done to compare performance of agents adopting various models or strategies. In the literature, there is no similar decision-making model for comparison. So, we choose two of the existing reputation models: Sabater and Sierra's REGRET Model [SS01] and Mui *et al.*'s Computational Model of trust and reputation [MMH02] for comparison. We have not implemented Marsh's model [Mar94] because agents' competence is irrelevant in our model, and the way trust should be modified is not defined, as discussed in Chapter 5. Also, Mui *et al.*'s Reputation Tic-for-tat [MHM02] is not implemented, either. This is because it cannot handle the case when a receiver receives more than one message at a time. Mui *et al.*'s another model [MMH02] is used for comparison instead.

6.1 Simulation Settings

Simulations are done to compare performance of agents employing our Trust/Honesty Model with performance of agents adopting other models or strategies. The setting of the simulation is as follows. We include receivers and senders adopting our Trust/Honesty Model. In addition, we include receivers and senders adopting other models or strategies. For non-adaptive receiver adopting our Trust/Honesty Model, a negative stubbornness and a risk attitude of 0.2 are used. For adaptive receiver adopting our Trust/Honesty Model, randomly generated stubbornness and risk attitude are used. For receivers adopting Sabater and Sierra's REGRET Model [SS01] and Mui *et al.*'s Computational Model of trust and reputation [MMH02], reputation and trust are calculated with the parameters suggested in these papers. These two receivers choose to believe the message from a sender with the maximum reputation when several messages are received at a time. If only one message is received, receiver adopting Mui *et al.*'s Computational Model chooses to believe the message if the sender's reputation is greater than 0.5, as suggested in the paper [MHM02]. Since Sabater and Sierra have not suggest any threshold and 0.5 is general enough to be a threshold, receiver adopting Sabater and Sierra's REGRET Model also chooses to believe the message if the sender's reputation is greater than 0.5 when only one message is received. Receivers adopting the "Choose Maximum Reputation" strategy chooses to believe the message from a sender with the maximum reputation when several messages are received at a time, where the reputation is calculated as suggested in Chapter 3. Using this calculation, an agent may have negative reputation. So, if only one message is received, this receiver chooses to believe the message if and only if the reputation of the message sender is positive. Similarly, receiver adopting the "Choose Maximum Utility" strategy choose the message with maximum utility to believe when several messages are received at a time, and chooses to believe the message if the utility of the message is greater than 0.5, where the utility is normalized to 1. Finally, receivers adopting the Random strategy randomly choose to believe a message when several messages are received at a time, and randomly choose to believe or not to believe the message when only one message is received.

For non-adaptive sender adopting our Trust/Honesty Model, a risk attitude of 0.4 and sincerity of 0.8 are used. Again, sincerity and risk attitude are randomly generated. Senders adopting the REGRET Model and Mui *et al.*'s model choose to tell lies if the target receiver has good reputation of being deceivable. The sender adopting the 100% Truth strategy always tells the truth. The sender adopting the Random 50% Truth strategy randomly tells 50% of truth and the one adopting the

Always Lie strategy always tells lies.

In each round, a random semi-competitive scenario is virtually generated. Each sender decides whether to tell a lie to a receiver according to its adopted strategy. Therefore, it is possible that a receiver receives more than one message at a time. Each receiver then chooses whether to believe the message according to its adopted strategy. Note that a receiver adopting the Trust/Honesty Model may believe no message at all if the persuasiveness of the messages it receives are all less than its stubbornness. At the end of each round, a receiver gains if it has believed a true message, or loses if it has believed a lie. On the other hand, a sender gains if the receiver has believed its message, or loses if the receiver has not. Then all agents update the impressions, reputations, trustworthiness, and deceivability accordingly. 1,000 games are simulated and each game contains 5,000 rounds.

6.2 Performance in Semi-competitive Environment

6.2.1 Performance of Receivers

Fig. 6.1 shows the average utility gain[1] of receivers adopting various models or strategies when interact with a specific type of senders, where mixed population means the population of senders contains senders with different models and strategies. In the figure, maximum possible utility means the maximum utility a receiver can possibly gain if it is so smart as to always choose the right message to believe, and has never been cheated. Note that this just serves as a benchmark for the comparison.

From the simulation results, adaptive receiver adopting our Trust/Honesty Model have the best performance in general. In particular, when interact with adaptive senders adopting our Trust/Honesty Model, the adaptive receiver obtains 84% of the maximum possible utility and outperforms receivers adopting other existing models

[1] Rounded up to the nearest integer.

Receiver's models/strategies	Average utility gain of receivers when interact with senders adopting							
	Non-adaptive Trust/Honesty Model	Adaptive Trust/Honesty Model	Sabater & Sierra's REGRET Model [SS01]	Mui *et al.*'s Computational Model [MMH02]	100% Truth strategy	50% Truth strategy	Always Lie strategy	Mixed population
Maximum possible utility	1980	1980	1980	1980	1980	1980	1980	1980
Non-adaptive Trust/Honesty Model	1092	848	803	817	1347	489	-34	902
Adaptive Trust/Honesty Model	1089	1662	1001	1017	1357	633	-4	1499
Sabater and Sierra's REGRET Model [SS01]	558	256	984	962	1346	448	-29	506
Mui et al.'s Computational Model [MMH02]	548	315	956	999	1354	486	-32	472
Choose Maximum Reputation	573	386	875	944	1254	487	-28	493
Choose Maximum Utility	-909	-835	-627	-753	624	-1036	-583	-834
Random	-851	-735	-946	-834	557	-878	-862	-727

Figure 6.1: Performance of receivers in semi-competitive environment

or strategies by 2 to 5 times. In another case when interact with senders adopting Always Lie strategy, the adaptive receiver outperforms other receivers by at least 7 times. When interact with a mixed population of all types of senders, the adaptive receiver obtains 76% of the maximum possible utility and outperforms others by about 3 times. Although when interact with non-adaptive senders adopting our Trust/Honesty Model, utility of the adaptive receiver is only comparable to that of the non-adaptive receiver adopting our Trust/Honesty Model, it outperforms other receivers by at least 2 times.

There are several reasons for the outstanding performance of the adaptive receiver adopting our Trust/Honesty Model. First, the adaptive receiver can learn to become more risk-seeking and less stubborn when the sender brings benefits to it, and it can learn to become more risk-averse and stubborn when the sender lies, while other receivers cannot. The adaptive receiver outperforms other receivers especially when interact with the adaptive senders. When the adaptive senders tell a certain number of truths, their reputation increase and receivers will believe them. Then the adaptive senders will become more risk-seeking and less sincere and tell more lies, but the adaptive receiver will be more risk-averse and stubborn after it is cheated and stop believing the adaptive sender if it continues to lie. However, as the senders' reputation are still high, receivers with other reputation models or reputation-related strategies will still believe them. Although the reputation of the adaptive senders will drop and these receivers will not believe the adaptive senders for some rounds, the adaptive senders will learn to be more risk-averse and more sincere and tell more truth to increase their reputation again. When their reputation are high, these receivers are cheated again. However, as the adaptive receiver will be more risk-averse and stubborn after it is cheated and stop believing the adaptive sender if it continues to lie, the adaptive senders will learn to be honest to the adaptive receiver. As a result, the adaptive receiver outperforms receivers adopting Sabater and Sierra's REGRET Model, Mui *et al.*'s Computational Model and the Choose Maximum Reputation strategy very significantly. In an extreme case when interact with senders adopting Always Lie strategy, the adaptive receiver does not lose much as it can learn to stop believing such lair very quickly. Although the adaptive receiver will increase its risk attitude and stubbornness when it has not believe any message for a number of rounds, as the receiver has not gain any utility, the percentage increase will be too small to be significant.

Second, receivers adopting Sabater and Sierra's REGRET Model, Mui *et al.*'s Computational Model and the Choose Maximum Reputation strategy only consider reputation in making decisions, but do not take into account the utility that will be brought by the messages. As a result, these receivers choose to believe those messages from senders with good reputation but the messages may bring a very low utility. Also, they miss some chances to earn from the less reputed senders. On the other hand, the adaptive and the non-adaptive receivers with our Trust/Honesty Model make a balance on trustworthiness of senders and utility of the messages, so these two receivers earn more in general. In addition, the adaptive percentage is proportional to the utility loss. This means that the more an adaptive receiver has lost, the faster it adapts and the faster it stops believing the lies.

When interact with senders adopting 100% Truth strategy, receivers adopting models or strategies related to reputation have high utilities and their performance are similar. This is because the senders always tell the truth and have high reputation, which always gain receivers' trust. However, the utilities of receivers adopting the Choose Maximum Utility strategy and Random strategy are not as good as the others'. This is because they do not believe all the messages they received. In other cases, receivers adopting the Choose Maximum Utility strategy and the Random strategy end up with negative utilities, because they are easily cheated. In all cases, performance of receivers adopting Sabater and Sierra's REGRET Model, Mui *et al.*'s Computational Model, and the Choose Maximum Reputation strategy are similar. This is because all of the three strategies choose the sender with maximum reputation to believe, only the ways they calculate the reputation are different.

6.2.2 Performance of Senders

Fig. 6.2 shows the average utility gain[2] of senders adopting various models or strategies when interact with a specific type of receiver, as well as a mixed population of receivers with different models and strategies. Again, maximum possible utility means the maximum utility that a sender can possibly get if it can always gain receivers' trust. Again, this serves only as a benchmark for comparison.

From the simulation results, the adaptive sender adopting our Trust/Honesty Model outperforms other senders in general. Only the sender using 100% Truth strategy has a comparable performance. However, sender adopting the 100% truth strategy cannot get the highest utility because the receivers with our Trust/Honesty Model may not believe its message if the utility is low. In particular, the adaptive sender outperforms other senders by 2 to 5 times when interacting with adaptive

[2]Rounded up to the nearest integer.

Sender's models/strategies	Average utility gain of senders when interact with receivers adopting							
	Non-adaptive Trust/Honesty Model	Adaptive Trust/Honesty Model	Sabater & Sierra's REGRET Model [SS01]	Mui *et al.*'s Computational Model [MMH02]	Choose Maximum Reputation	Choose Maximum Utility	Random	Mixed population
Maximum possible utility	1868	1868	1868	1868	1868	1868	1868	1868
Non-adaptive Trust/Honesty Model	998	768	777	786	854	956	712	923
Adaptive Trust/Honesty Model	996	1620	1010	1003	982	976	713	1485
Sabater and Sierra's REGRET Model [SS01]	696	328	922	968	913	967	709	475
Mui et al.'s Computational Model [MMH02]	709	397	904	983	933	949	711	482
100% Truth strategy	714	1068	956	981	940	944	713	726
50% Truth strategy	-609	-1360	-1123	-1268	-1373	951	710	-620
Always Lie strategy	-1815	-2387	-1799	-1747	-1848	980	711	-1903

Figure 6.2: Performance of senders in semi-competitive environment

receivers adopting our Trust/Honesty Model. Also, the adaptive sender outperforms other senders by about 3 times when interacting with a mixed population of all types of receivers.

When interacting with receivers adopting Choose Maximum Utility and Random strategies, all senders have similar performance and senders with Random 50% Truth and Always Lie strategies have positive utility gain because these two strategies are independent of reputation. In other cases, senders using Random 50% Truth strategy and Always Lie strategy have negative utilities because they are hardly believed by the receivers as they have low reputation. In all cases, senders adopting Sabater and Sierra's REGRET Model and Mui *et al.*'s Computational Model have similar performance. This is because both of the two senders choose to cheat the receiver with maximum reputation of deceivable, only the ways they calculate the reputation are different. The performance of these two senders are not as good as that of the senders adopting our Trust/Honesty Model. This is because these agents only consider reputation in making decisions, but do not take into account the utility, which means they tell lies if the receiver has good reputation of being deceivable, even if utility gain is small. As a result, these senders tell more lies than senders with our Trust/Honesty Model, which make a balance on deceivability of receivers and utility of lying. So, if the utility of lying is not good enough, senders with our Trust/Honesty Model will choose not to lie, which maintain their reputation.

Another reason for the outstanding performance of the adaptive sender adopting our Trust/Honesty Model is that it can learn to be more risk-seeking and less sincere when the receiver believes its messages and learn to be more risk-averse and sincere when it lose the receiver's trust, while other senders cannot adapt. In particular, when interact with the adaptive receivers, utilities of senders adopting Sabater and Sierra's REGRET Model and Mui *et al.*'s Computational Model are only one-forth of that of the adaptive sender. This is because when the adaptive receivers believe their messages, the receivers have good reputation of deceivable, then they cheat the receivers. However, the receivers adapt and do not believe their messages. These senders tell the truth again only when the receivers' reputation drop. On the other hand, the adaptive sender learns quickly as the adaptive percentage is proportional to utility lost. As described in section 6.2.1, the adaptive sender learns to be honest to the adaptive receivers and gains the receivers' trust.

6.3 Performance when Interact with Strategic Senders

In this section, simulations are done to compare the performance of agents adopting various models or strategies when interact with strategic senders. The first strategic sender tells a lie bringing a loss of 1 in every ten rounds, and tells the truth but brings only a utility gain of 0.01 in the rest of the time. The second strategic sender tells a lie bringing a loss of 0.01 in every ten rounds, and tells the truth but brings a utility gain of 1 in the rest of the time. The third strategic sender tells a lie bringing a loss of 0.01 in every ten rounds, and tells the truth but brings a utility gain of 0.01 in the rest of the time. Strategic sender 4 tells a lie bringing a loss of 1 in every

Receiver's models/strategies	Average utility gain of receivers when interact with				
	Strategic Sender 1	Strategic Sender 2	Strategic Sender 3	Strategic Sender 4	Strategic Sender 5
Maximum possible utility	45	4500	45	4500	500
Non-adaptive Trust/Honesty Model	-24	4494	39	3999	349
Adaptive Trust/Honesty Model	41	4495	40	4000	497
Sabater and Sierra's REGRET Model [SS01]	-454	4493	39	3998	0
Mui et al.'s Computational Model [MMH02]	-455	4494	38	3998	0
Choose Maximum Reputation	-452	4492	37	3996	0
Choose Maximum Utility	-500	4495	0	4000	500
Random	-228	2245	20	2000	228

Figure 6.3: Performance of receivers when interact with strategic senders

ten rounds, and tells the truth but brings a utility gain of 1 in the rest of the time. Strategic sender 5 tells a truth bringing a gain of 1 in every ten rounds, and tell lies bringing a utility loss of 0.01 in the rest of the time. The results are shown in Fig. 6.3.

6.3.1 Senders Telling More Truths than Lies

When interacting with the first strategic sender, which tells a lie bringing a loss of 1 in every ten rounds, and tells the truth but brings only a utility gain of 0.01 in the rest of the time, adaptive receiver adopting our Trust/Honesty Model has the best performance. When this receiver is cheated by the sender, it becomes more risk-averse. This means that it is more cautious for messages with high utilities. As a result, after being cheated for a few times, it does not believe the messages with high utilities anymore. On the other hand, it chooses to believe the messages with low utilities when the persuasiveness of the message is higher than its stubbornness. Although, it increases its risk attitude and lowers its stubbornness when it has believed the right messages for a number of rounds, the adaptive percentage is not significant as utility gain is little and utility lost is relatively large in comparison. In this way, adaptive receiver adopting our Trust/Honesty Model obtains a very outstanding performance approaching the maximum possible utility. For the non-adaptive receiver adopting our Trust/Honesty Model, as its risk attitude and stubbornness do not change during the game, it only stops believing the lie when the trustworthiness of the lair becomes low. So, its utility is negative. However, its performance is much better than performance of agents adopting other models or strategies. Agents adopting Sabater and Sierra's REGRET Model, Mui *et al.*'s Computational Model and the Choose Maximum Reputation strategy have similar results. All of these three agents do not take utility into account in making decision. In addition, agents adopting Sabater and Sierra's REGRET Model and Mui *et al.*'s Computational Model consider a sender to have good reputation only because the sender tell more truths than lies. As the strategic sender tells much more truth than lies, these two agents will always believe this strategic sender. Since the lies bring great loss in utility, the performance of these agents are very low. Agent adopting the Choose Maximum Utility strategy believes all the lies but not any truth as the utilities of the truths are very low and the utilities of the lies are very high, so this agent also has very bad performance. Agent adopting a Random strategy randomly believes about half of the message, it also has a negative utility gain.

When interacting with strategic senders 2 to 4, all agents with reputation-related models and strategies obtain very high utilities approaching the maximum possible utility. This is because these senders tell the truths most of the time and the total lost brought by the lies are low in comparison. As a result, agents with our Trust/Honesty Model always believe this sender due to its high trustworthiness. Although the adaptive agent lowers its risk attitude and increases its stubbornness when it is cheated, it still always believe these senders. This is because utility gain is much more than lost, which makes this adaptive percentage too low to be significant. Agents with Sabater and Sierra's REGRET Model, Mui *et al.*'s Computational Model and the Choose Maximum Reputation strategy always believe these senders since the senders tell the truths most of the time. Agent using the Choose Maximum Utility strategy always believes the truths if the truths bring high utilities, believes no messages if all messages bring very low utilities, and believes all messages if all messages bring high utilities. Agent using the Random strategy only chooses about half of the message to believe, so it misses some chances to earn which makes it has a relatively low utility gain.

6.3.2 Senders Telling More Lies than Truths

It is easy to understand that if the sender tells much more lies than truths, receivers with reputation-related models or strategies will not believe the sender because of the low reputation of the sender. Let us consider strategic sender 5, which tells a truth bringing a gain of 1 in every ten rounds, and tell lies bringing a utility loss of 0.01 in the rest of the time. Receiver using Choose Maximum Utility strategy can get a maximum utility because it only chooses to believe the messages with high utilities, which are the truths. It does not believe any lies as the utilities of the lies are too small. This is just a special case in which receiver using the Choose Maximum Utility strategy performs well. In other simulations shown before, receiver using this strategy generally perform not very well. Adaptive receiver with our Trust/Honesty Model can get a utility very close to the maximum utility. This receiver and the non-adaptive one outperform receivers with other reputation-related models and strategies. This is because reputation of the liar is low and utilities of the lies are also low, which make the adaptive receiver do not believe the lies. As the adaptive receiver has not believed any message for a number of rounds, it will increase its risk attitude. As it become risk-seeking enough, it will believe messages with high utilities, which are the truths. Although the non-adaptive receiver does not perform as good as the adaptive one, its utility gain is much better than other receivers with reputation-related models or strategies. This is because it also consider utility in making decisions. On the other hand, receivers adopting Sabater and Sierra's REGRET Model, Mui *et al.*'s Computational Model and Choose Maximum Reputation strategy consider only reputation in decision making. As this sender tell lies most of the time, it has a very low reputation in view of these three receivers, so these receivers do not believe this sender at all and thus cannot gain anything. Again, receiver using Random strategy gets about half of the maximum possible utility.

6.4 Summary

These experiments show that our Adaptive Trust/Honesty Model significantly outperforms other trust models. This is because the adaptive strategies enable agents to learn from their experiences and change their parameters accordingly. Another reason is that the Trust/Honesty Model makes a balance on reputation and expected utility in making decisions, while models or strategies considering only reputation or only expected utility cannot achieve high utility.

Chapter 7

Conclusion and Future Work

7.1 Conclusions

This thesis describes methods for agents to choose their actions and communication actions in semi-competitive environments. Semi-competitive environment is an environment in which cooperation and competition can both exist.

For agents to represent their knowledge, predict other agents' actions and choose their own actions. Gmytrasiewicz and Durfee propose the Recursive Modeling Method (RMM). RMM is recursive as it not only represents an agent's own preferences, abilities and beliefs about the world, but also represents the beliefs the agent has about other agents, the beliefs it has about other agents' beliefs, and so on, forming an infinite hierarchy. However, the authors have made an assumption that the belief hierarchy is finite and terminates the hierarchy explicitly by a probabilistic approximation at the point where an agent has no sufficient information to model other agents. We improve the original design of RMM by Recursive Formulas, with which no assumption and approximation are made.

For agents to choose their communication actions, we introduce our Trust/Honesty Model. In semi-competitive environments, agents have incentives to be honest as well as dishonest. So, being a sender, agent needs to choose whether to tell lies or to tell the truth. On the other hand, being a receiver, agent needs to choose whether to believe the received message or which message to believe. From a receiver's point of view, we introduce a Trust Model, which enables the receiver to determine whether to trust the received messages. In the Trust Model, we first differentiate and define the terms impression, reputation and trustworthiness. We introduce how a receiver decides on which message to believe and follow by comparing the persuasiveness of the message with its stubbornness to the sender, where persuasiveness of a message is calculated from the trustworthiness of the message sender, utility brought by the message and risk attitude of the receiver. From a sender agent's point of view, we introduce an Honesty Model, which enables the sender to determine whether to be honest. To do so, we propose to calculate the temptation of lying from deceivability of the receiver, utility of lying and risk attitude of the sender, and compare it with the sender's sincerity to the receiver.

Furthermore, we improve our Trust/Honesty Model by an adaptive strategy. With the adaptive strategy, receiver can learn to be more risk-averse and stubborn after it is cheated and learn to be more risk-seeking and less stubborn to senders who bring benefits. As symmetry, an adaptive sender can learn to be more risk-averse and sincere after it lost the receiver's trust and learn to be more risk-seeking and less sincere to deceivable receivers. We relate the adaptive percentage to the utility that an agent has gained and has lost. This is because the more the receiver lose, the more risk-aver and the more stubborn it will become. This mimics the model in human interaction.

Simulations shown that agents with our Adaptive Trust/Honesty Model perform much better than agents with other existing models or strategies. This is because existing models only take reputation into account in making decision, but agents with our Trust/Honesty Model consider both trustworthiness/deceivability of the opponents and utility brought by the messages. In addition, agents with the adaptive strategy can adapt to the environment, learn from experiences and prevent themselves from being cheated or learn to be more honest.

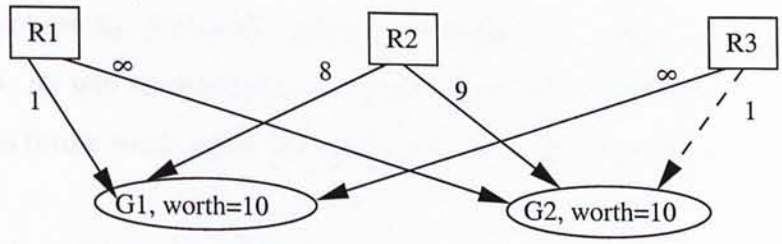

Figure 7.1: Example of agent manipulation

7.2 Future Work

7.2.1 Agent Manipulation

There is another approach to perform cheating. Let us consider the following example. There are two goals G_1 and G_2, both with worth 10. G_1's priority ordering of agents is $R_2 > R_1 > R_3$ and G_2's priority ordering of agents is $R_2 > R_3 > R_1$. The cost for R_1 to obtain G_1 is 1, the costs for R_2 to obtain G_1 and G_2 are 8 and 9 respectively, and the cost for R_3 to obtain G_2 is 1. R_1 is inaccessible to G_2 and R_3 is inaccessible to G_1. Initially, R_1 and R_2 know the presence of G_1 and G_2 but R_3 knows nothing. The example setting is shown in Fig. 7.1. By applying RMM, it can be seen that R_1 can get maximum payoff by obtaining G_1, provided that R_2 does not obtain G_1. As R_2 has the highest priority in G_1's priority ordering of agents and R_2 can get maximum payoff by obtaining G_1, the only way for R_1 to prevent R_2 from obtaining G_1 is to lie to R_2 that there is another goal, which gives R_2 a higher payoff than G_1 does. As R_1 knows that telling a lie to R_2 will decrease its trustworthiness, so instead of telling the lie by itself, R_1 tells R_3 the information about G_2 and manipulate R_3 to tell the lie. If R_3 believe the information provided by R_1, it will know that it can get a maximum payoff by obtaining G_2, provided that R_2 does not obtain G_2. From R_3's point of view, it knows only about G_2, R_2 has the highest priority in G_2's priority ordering of agents and R_2 can get maximum payoff by obtaining G_2, so R_3 thinks the only way for it to prevent R_2 from obtaining the worth of G_2 is to tell R_2 that there is another goal, which can give R_2 a higher payoff than G_2 does. As R_3 has no other information, the only way for it to get any worth is to cheat R_2. Follow the utility-maximizing behavior, R_3 will do so, and if it succeeds, R_2 will be directed to obtain the fake goal, which is what R_1 wants to achieve. As future work, agent manipulation can be investigated.

7.2.2 Algorithm for Solving Recursive Formulas

We have not mentioned how the recursive formulas can be solved in a computationally tractable way, as it is out of the scope of this thesis. As future work, an algorithm for solving the recursive formulas can be designed. In addition, one can do simulations and compare the performance of the original RMM with that of the improved RMM.

7.2.3 Fuzzy Trust/Honesty Model

The concepts in the Trust/Honesty Model are rather fuzzy. An agent can have a *good* impression on one agent, or have a *bad* impression on another agent, where *good* and *bad* are fuzzy terms. Also, risk averse, sincere, neutral stubborn, ..., are all fuzzy terms. So, as future work, a fuzzy Trust/Honesty Model can be developed.

7.2.4 Opinion from the Mass

In the proposed Trust/Honesty Model, agents make rational decision by using impression, reputation, risk attitude, and expected payoff. An agent chooses to believe a message if persuassiveness of the message is greater than the agent's stubbornness. However, in human communication model, opinion from the mass can also affect the decision. For example, a person receives a message saying the stock market will rise and this person chooses to believe the message because of the trustworthy source and high expected payoff. However, ten other people, who also receive the same message, choose not to believe the message. Then the person who chooses to believe the message before may consider not believing the message due to other people's decision. As future work, elements can be added to the model to handle this case.

7.2.5 Network Application

There is a fundamental assumption in the current generation of ad hoc networks, which is that the nodes will cooperate and will not cheat [DD03]. However, such assumption may no longer be valid if the nodes in the network do not have a common goal. The environment will then become semi-competitive. As future work, our Trust/Honesty Model can be applied in an ad hoc network.

Bibliography

[Axe84] R. Axelrod. *The Evolution of Cooperation*. New York: Basic Books, 1984.

[Bra92] A. Brandenburger. Knowledge and equilibrium in games. *Journal of Economic Perspectives*, 6:83–101, 1992.

[CF98] C. Castelfranchi and R. Falcone. Principles of trust for mas: Cognitive anatomy, social importance, and quantification. In *Proceedings of the Third International Conference on Multiagent Systems*, pages 72–79, 1998.

[CFP03] C. Castelfranchi, R. Falcone, and G. Pezzulo. Trust in information sources as a source for trust: A fuzzy approac. In *Proceedings of The Second International Joint Conference on Autonomous Agent and Multiagent Systems*, pages 89–96, 2003.

[DD03] P. Dewan and P. Dasgupta. Trusting routers and relays in ad hoc networks. In *Proceedings of The 2003 International Conference on Parallel Processing Workshops*, 2003.

[FC01] R. Falcone and C. Castelfranchi. Social trust: A cognitive approach. In *Trust and Deception in Virtual Societies*, pages 55–90. Kluwer Academic Publishers, 2001.

[Fis81] P. Fishburn. *The foundations of expected utility*. D. Reidel Publishing Company, 1981.

[GD92] P. J. Gmytrasiewicz and E. H. Durfee. Decision-theoretic recursive modeling and the coordinated attack problem. In *Proceedings of the First International Conference on AI Planning Systems*, pages 88–95, 1992.

[GD93] P. J. Gmytrasiewicz and E. H. Durfee. Toward a theory of honesty and trust among communicating autonomous agents. *Group Decision and Negotiation*, 2:237–258, 1993.

[GD95] P. J. Gmytrasiewicz and E. H. Durfee. A rigorous, operational formalization of recursive modeling. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 125–132, 1995.

[GD00] P. J. Gmytrasiewicz and E. H. Durfee. Rational coordination in multi-agent environments. *Autonomous Agents and Multi-Agent Systems*, 3(4):319–350, 2000.

[GD01] P. J. Gmytrasiewicz and E. H. Durfee. Rational communication in multi-agent environments. *Autonomous Agents and Multi-Agent Systems*, 4:233–272, 2001.

[GDW91] P. J. Gmytrasiewicz, E. H. Durfee, and D. Wehe. A decision-theoretic approach to coordinating multiagent interactions. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, pages 62–68, 1991.

[GG00] A. Glass and B. Grosz. Socially conscious decision-making. In *Proceedings of the Fourth International Conference on Autonomous Agents*, pages 217–224, 2000.

[GL03] N. Griffiths and M. Luck. Coalition formation through motivation and trust. In *Proceedings of The Second International Joint Conference on Autonomous Agent and Multiagent Systems*, pages 17–24, 2003.

[htta] http://dictionary.cambridge.org/. Cambridge dictionaries online.

[httb] http://www.webster.com/. Merriam-webster online.

[LJS+03] X. Luo, N. R. Jennings, N. Shadbolt, H. F. Leung, and J. H. M. Lee. Fuzzy constraint based model for bilateral, multi-issue negotiations in semi-competitive environments. *Artificial Intelligence*, 148:53–102, 2003.

[Mar94] S. Marsh. *Formalising Trust as a Computational Concept*. PhD thesis, University of Stirling, 1994.

[MHM02] L. Mui, A. Halberstadt, and M. Mohtashemi. Notions of reputation in multi-agent systems: A review. In *Proceedings of Autonomous Agents and Multi-Agent Systems*, 2002.

[MMA+01] L. Mui, M. Mohtashemi, C. Ang, P. Szolovits, and A. Halberstadt. Ratings in distributed systems: A bayesian approach. In *Workshop on Information Technologies and Systems*, 2001.

[MMH02] L. Mui, M. Mohtashemi, and A. Halberstadt. A computational model of trust and reputation. In *Proceedings of 35th Hawaii International Conference on System Science*, 2002.

[Rai82] H. Raiffa. *The art and science of negotiation*. Harvard University Press, 1982.

[RG85] J. S. Rosenschein and M. R. Genesereth. Deals among rational agents. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 91–99, 1985.

[RLM01] J. C. Rubiera, J. M. M. Lopez, and J. D. Muro. A fuzzy model of reputation in multi-agent systems. In *Proceedings of the Fifth International Conference on Autonomous Agents*, pages 25–26, 2001.

[SS01] J. Sabater and C. Sierra. Regret: A reputation model for gregarious societies. In *Proceedings of Fourth International Workshop on Deception, Fraud and Trust in Agent Societies*, 2001.

[SW00] S. Sen and G. Weiss. Learning in multiagent systems. In *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, pages 259–298. The MIT Press, 2000.

[WJ95] M. Wooldridge and N. R. Jennings. Intelligent agents: Theory and pratice. *The Knowledge Engineering Review*, 10(2):115–152, 1995.

[YS01] B. Yu and M. P. Singh. Towards a probabilistic model of distributed reputation management. In *Proceedings of Fourth International Workshop on Deception, Fraud and Trust in Agent Societies*, pages 125–137, 2001.

[YS03] B. Yu and M. P. Singh. Detecting deception in reputation management. In *Proceedings of The Second International Joint Conference on Autonomous Agent and Multiagent Systems*, pages 73–80, 2003.

[ZR89] G. Zlotkin and J. S. Rosenschein. Negotiation and task sharing among autonomous agents in cooperative domains. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, pages 912–917, 1989.

[ZR90] G. Zlotkin and J. S. Rosenschein. Negotiation and conflict resolution in non-cooperative domains. In *Proceedings of the National Conference on Artificial Intelligence*, pages 100–105, 1990.

Publications

Full Length Conference Papers

- K.M. Lam and H.F. Leung. An Infinite Belief Hierarchy based on the Recursive Modeling Method. In Proceedings of Sixth Pacific Rim International Workshop on Multi-Agents, pages 25–36, 2003.
- K.M. Lam and H.F. Leung. Honest, Trust, and Rational Communication in Multi-agent Semi-Competitive Environments. In Lecture Notes in Computer Science, volume 2891, pages 158–169, 2003. (Runner-Up paper award in Sixth Pacific Rim International Workshop on Multi-Agents)
- K.M. Lam and H.F. Leung. A Trust/Honesty Model in Multi-agent Semi-competitive Environments. Seventh Pacific Rim International Workshop on Multi-Agents, 2004.

Conference Poster Papers

- K.M. Lam and H.F. Leung. Rational Communication in Multi-agent Semi-Competitive Environments. In Proceedings of The Second International Joint Conference on Autonomous Agents and Multiagent Systems, pages 1044–1045, 2003.

Submitted Papers

- K.M. Lam and H.F. Leung. An Adaptive Strategy for Trust/Honesty Model in Multi-agent Semi-competitive Environments. The 16th IEEE International Conference on Tools with Artificial Intelligence, 2004.
- K.M. Lam and H.F. Leung. A Trust/Honesty Model, with Adaptive Strategy, for Multiagent Semi-competitive Environments. Journal Autonomous Agents and Multi-Agent Systems.