Preference Uncertainty and Trust in Decision Making

by

Mubarak AL-Mutairi

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Systems Design Engineering

Waterloo, Ontario, Canada, 2007

Author’s Declaration for Electronic Submission of a Thesis

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Mubarak AL-Mutairi

Abstract

A fuzzy approach for handling uncertain preferences is developed within the paradigm of the Graph Model for Conflict Resolution and new advances in trust modeling and assessment are put forward for permitting decision makers (DMs) to decide with whom to cooperate and trust in order to move from a potential resolution to a more preferred one that is not attainable on an individual basis. The applicability and the usefulness of the fuzzy preference and trust research for giving an enhanced strategic understanding about a dispute and its possible resolution are demonstrated by employing a realworld environmental conflict as well as two generic games that represent a wide range of real life encounters dealing with trust and cooperation dilemmas.

The introduction of the uncertain preference representation extends the applicability of the Graph Model for Conflict Resolution to handle conflicts with missing or incomplete preference information. Assessing the presence of trust will help to compensate for the missing information and bridge the gap between a desired outcome and a feared betrayal. These advances in the areas of uncertain preferences and trust have potential applications in engineering decision making, electronic commerce, multiagent systems, international trade and many other areas where conflict is present.

In order to model a conflict, it is assumed that the decision makers, options, and the preferences of the decision makers over possible states are known. However, it is often the case that the preferences are not known for certain. This could be due to lack of information, impreciseness, or misinformation intentionally supplied by a competitor. Fuzzy logic is applied to handle this type of information. In particular, it allows a decision maker to express preferences using linguistic terms rather than exact values. It also makes use of data intervals rather than crisp values which could accommodate minor shifts in values without drastically changing the overall results. The four solution concepts of Nash, general metarationality, symmetric metarationality, and sequential stability for determining stability and potential resolutions to a conflict, are extended to accommodate the new fuzzy preference representation. The newly proposed solution concepts are designed to work for two and more than two decision maker cases. Hypothetical and real life conflicts are used to demonstrate the applicability of this newly proposed procedure.

Upon reaching a conflict resolution, it might be in the best interests of some of the decision makers to cooperate and form a coalition to move from the current resolution to a better one that is not achievable on an individual basis. This may require moving to an intermediate state or states which may be less preferred by some of the coalition members while being more preferred by others compared to the original or the final state. When the move is irreversible, which is the case in most real life situations, this requires the existence of a minimum level of trust to remove any fears of betrayal. The development of trust modeling and assessment techniques, allows decision makers to decide with whom to cooperate and trust. Illustrative examples are developed to show how this modeling works in practice.

The new theoretical developments presented in this research enhance the applicability of the Graph Model for Conflict Resolution. The proposed trust modeling allows a reasonable way of analyzing and predicting the formation of coalitions in conflict analysis and cooperative game theory. It also opens doors for further research

and developments in trust modeling in areas such as electronic commerce and multiagent systems.

Acknowledgements

Getting a doctoral degree is just like having a passport to enter a new phase of your life, in which you have to decide what to do next. A PhD is all about widening your academic horizon and committing yourself to learning. During this long journey, so many people became involved in making it worth the effort and thoroughly enjoyable. I am really grateful to my supervisors, Professor Keith W. Hipel and Professor Mohamed S. Kamel. Their enthusiasm, support, guidance, and companionship enabled me to achieve what I dreamed of. Their encouragement to adopt a system design approach to tackling my research problems and their continuous guidance were vital to my doctoral studies.

Professor Hipel has always been a great mentor and a source of motivation. Above all, his wisdom, perseverance, and enthusiasm often inspired me. His teaching skills and sensitivity to students' needs were of a great inspiration to me. I owe it for him winning the Distinguished Teaching Award by a Registered Student. He was a great model as a researcher, scientist, and faculty member.

Equally, I am so grateful to Professor Kamel for his enlightening discussions, moral support, and continuous advice and guidance. His extensive knowledge and experience, open mindedness, sharpness, sound judgment, and patient approach to research were very crucial to me in successfully completing this thesis. His ideas were particularly valuable in solving many of the challenges that I faced during the development of my ideas. I continued to be impressed by his articulate style of writing and expression. I would like to thank him for his mentoring, support, and encouragement.

Special thanks also go to Professor Gordon Huang, the External Examiner from the University of Regina, Professor Fakhri Karray, Professor Yang Wang, and Professor Khalid Hassanein, members of my thesis examination committee, for devoting valuable time from their busy schedules.

There are hardly words enough to thank my parents and family for all they have done for me. I am very grateful for their constant support and trust in me and all my endeavors. They didn’t loose patience with me even in my most challenging times. And, of course, thanks to my brothers Faihan and Ali who always brightened all my days and still do. THANK YOU!

Dedication

To my parents, Saad and Nuwayer.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Conflicts, Trust and Uncertainty

Conflicts were and will always be an integral part of human interactions. Conflicts are triggered by the decisions one makes and the actions one takes. Conflicts are usually ignited by a clash of interests (Fraser and Hipel, 1994, Hipel 2001). Those decisions could range from daily life decisions to decisions that could change history for decades. When American President, George W. Bush, decided to invade Iraq in the year 2003 against the advice of countries such as Canada and France, the most pessimistic member of the American administration did not predict the resulting catastrophic outcomes which are taking place in Iraq right now. This in part is due to lack of credible information and in part to the absence of trust between the American and the different local Iraqi groups.

In order to make realistic decisions in conflicting situations, formal conflict analysis tools and solution concepts are needed to "assist in the understanding, modeling, and analysis of conflict" (Fang et al., 1993). Conflict analysis techniques are known and widely accepted procedures for resolving conflicts arising in areas such as engineering, economics, and politics (Bennett, 1995; Binmore, 1992; Brams, 1994; Fang et al., 1993; Osborne, 2003).

It is important to realize that modeling, in general, is a mathematical representation that tries to capture the key characteristics of real life problems. Accurate representation may result in an added system complexity while too much simplicity may fail to capture some of the key components of the system or problem under study. A compromise between these two conflicting criteria is required to model, analyze, and predict a realistic resolution in accordance with the principle of Occam's Razor (Hipel and McLeod, 1994).

1.2 Traditional Approaches to Conflict Resolution and Coalition Analysis

In order to better comprehend, model and analyze conflicts, a number of methodologies have been proposed. Among, but not limited to, those methodologies, are the graph model for conflict resolution (Fang et al.,1993), conflict analysis (Fraser and Hipel, 1984), theory of moves (Brams, 1994), theory of fuzzy moves (Kandel and Zhang, 1998; Li et al., 2001), drama theory (Bennett and Howard, 1996; Bennett, 1998; Bryant, 2003; Howard, 1999; Howard et al., 1992; Howard, 1999; Howard, 1994), hypergame analysis (Bennett, 1977, 1980; Wang et al., 1988), and metagame analysis (Howard, 1971).

The common dominator among all of them is that they are based on game theoretic approaches. These approaches "have game-theoretic roots – all are essentially game theory variants that have been designed to yield better decision advice or more compelling structural insights." (Kilgour, 1995). The publication of the book entitled *The Theory of Games and Economic Behavior* (Von Neumann and Morgenstern, 1944) laid the foundations for the modern development of game theory. It was not until 1950 when John Nash introduced the concept of Nash stability (Nash, 1950, 1951) in game theory, for which he was awarded the Nobel Prize in economics in the year 1994, that game theory started to increase in importance.

In the modeling stage, the key decision makers (DMs), sometimes referred to as players, participants, or stakeholders, the options or course of actions available to each DM, and the relative preferences of each DM over the set of feasible states need to be identified. The analysis stage is a systematic strategic assessment of the conflict using the available information about the DMs, options, and preferences. A predicted resolution or equilibria refers to a state which is stable for all DMs. To answer the what-if questions and tests the model's response to changes in the model's parameters, sensitivity analyses are employed. Changes in DMs, feasible states, preferences, or any combination of the three could be used as a tool to check the robustness of the model. In most cases, the output of a robust model wouldn't change significantly from minor changes in the parameters. For the situation in which a minor change in the parameters results in a drastic shift in the analytical results, the analyst needs to revisit the model and adjust and fine-tune the underlying assumptions.

Conflicts can be represented as games or "a model of interacting DMs" (Osborne, 2003). These games can be classified as cooperative or non-cooperative games where the available courses of action for each DM is referred to as his or her strategy. The combination of the selected strategies by each player are referred to as outcomes or states. Among the simplest and most intuitive approaches for analyzing conflicts is the Graph Model for Conflict Resolution (Fang et al., 1993).

1.3 Stability Analysis

To be able to assess the stability of a specific state for a specific DM, the preferences of that DM over the set of feasible states or the available course of actions, need to be identified using cardinal values or ordinal preferences. Cardinal values (Fishburn, 1988) could be utilized using some utility values (Von Neumann and Morgenstern, 1953). When using cardinal values, one is actually using real numbers that reflect both the preference and the strength of the preference. On the other hand, using the ordinal preferences only reflects the preference but does not tell anything about the preference's strength. In both cases, when the decision maker is unable to express a preference due to lack of information or uncertainties, the states are said to be equally preferred. Transitivity is assumed in both cardinal payoffs and ordinal preferences. By that, one means that if a decision maker prefers state p to state q and state q to state r that implies that state p is preferred to state r by the same decision maker.

In situations where transitivity is not assumed, one can compare the states in a pairwise fashion. A decision maker may prefer state p to state q and state q to state r but at the same time prefer state r to state p. A decision maker may prefer one state over another or may not have a preference among some states. In this case, the states are said to be equally preferred (Fang et al., 1993).

There are some situations where the DM's preferences could be expressed in a vague or imprecise fashion due to a lack of information. Several attempts have been put forward to handle preference's uncertainties within the paradigm of conflict analysis. The information gap model is a way to bridge the uncertainty in preference information based on the known preferences (Ben-Haim and Hipel, 2002). Other attempts included specifying the minimum required preference information to ensure stable equilibria using robustness analysis (Sakakibara et al., 2002). Li et al. (2004) introduced a new uncertain binary relationship for expressing a DM's preference uncertainty in conflict models.

AL-Mutairi et al. (2006 a, b, and d), acknowledge preference uncertainty and try to handle it by using means of fuzzy logic. The definitions of Nash, GMR, SMR, and SEQ have been extended to accommodate this fuzziness in information for both two and multiple participants cases. This line of research constitutes one major component of this thesis.

State stability is assessed based on the behavioral patterns of the DM of interest depending on the solution concept under consideration. Different solution concepts being applied within the paradigm of the Graph Model for Conflict Resolution include Nash stability (Nash) (Nash, 1950, 1951), general metarationality (GMR) (Howard, 1971), symmetric metarationality (SMR) (Howard, 1971), sequential stability (SEQ) (Fraser and Hipel, 1979, 1984), limited move stability (L_h, $h > 1$) (Kilgour, 1985; Zagare, 1984; Fang et al., 1993), and non-myopic stability (NM) (Brams and Wittman, 1981; Kilgour, 1984).

Applying these solution concepts within the graph model structure enables one to consider different key factors such as foresight level and strategic risk. The decision support system GMCR II (Hipel, et al., 1997; Fang, et al., 2003a,b; Peng, 1999) implements these solution concepts within GMCR in an easy, efficient, and user-friendly way. Equipped with an analysis engine and interactive menus, GMCR II generates equilibrium states under different solution concepts.

Upon reaching a resolution or an equilibrium for the conflict under study, it might be of interest to two or more of the DMs involved in that conflict to form a coalition in order to reach a more preferred equilibrium. Some conditions are applicable for an allowable coalition. First, the move has to be from one equilibrium to another more preferred equilibrium for at least two or more of the DMs. Second, it has to be not reachable by a unilateral move by one of the DMs. Last, the new move by the coalition can't be sanctioned by a move of none coalition members (Fraser and Hipel, 1984).

The coalition move to a more preferred equilibrium may include moving to a transitional intermediate state. This intermediate state might be less preferred by some of the coalition members and at the same time more preferred by other members. This raises the question of what motivates the coalition members to meet their obligation and move the coalition to its intended final state. More importantly, what motivates this coalition in the first place? AL-Mutairi et al (2005b, 2006c), introduce a new methodology to assess trust using fuzzy logic in order to study the feasibility of forming a coalition. This new proposed methodology was expanded to include possible applications in multiagent systems, electronic transactions, and decision making in general. The new concept of trust is also introduced to bridge the gap between the known and unknown information in situations of coalitions and interactions with entities for the first time.

1.4 Thesis Outline

The objective of this thesis is to redefine the solution concepts of Nash, GMR, SMR, and SEQ stabilities in order to accommodate preference uncertainties using fuzzy logic. These new solution concepts are to be defined for both two and multiple decision maker situations. Also a comprehensive and extended survey of trust from different disciplines is carried out. This helps to examine the current trust models and analyze where and why they are different. A fuzzy logic based modeling of trust is also proposed for use in coalition and decision making situations.

The reminder of this thesis is organized as follows. Chapter 2 begins with an overview of the graph model for conflict resolution and the decision support systems GMCR I and its descendent GMCR II. In Chapter 3, the four solution concepts of Nash, GMR, SMR and SEQ, are redefined to accommodate the new fuzzy preference structure for two-DM cases. These solution concepts have been applied to the famous Prisoner's Dilemma to show its applicability. In Chapter 4, the concepts of group unilateral moves and group unilateral improvements are introduced. The four solution concepts are also redefined for the multiple participants cases. A real world conflict of an underground water contamination is presented for analysis in this chapter. Chapter 5 gives an extensive assessment of trust research from different disciplines along with an overview of the existing trust models. Chapter 6 proposes a new trust modeling approach based on fuzzy logic. Chapter 7 concludes with a summary for the thesis' original contributions and directions of future research extensions and work. Figure 1.1 summarizes the thesis outline and the research methodologies.

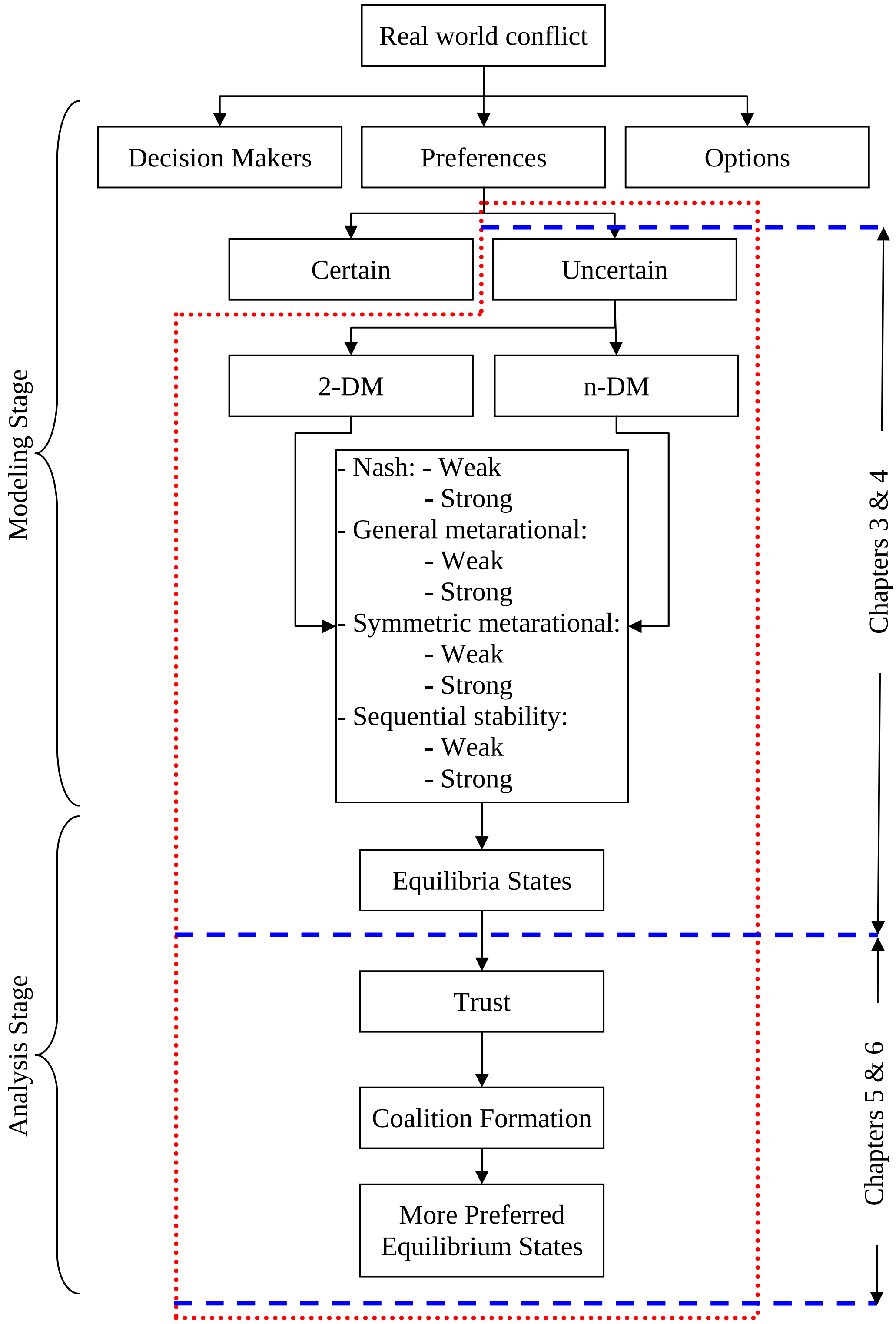

Figure 1.1: Thesis outline and research methodology.

Chapter 2

The Graph Model for Conflict Resolution

A variety of approaches have been put forward to better model and analyze real life conflicts. Most of the current modeling approaches have some game theoretical bases. This dates back to the year 1944 when Von Neumann and Morgenstern published their seminal book *Theory of Games and Economic Behavior*. In 1971, Howard introduced the metagame analysis (Howard, 1971) which was followed by the conflict analysis approach of Fraser and Hipel (1979, 1984). The most feasible and intuitive approach to conflict resolution is the Graph Model for Conflict Resolution, which was developed by Fang, Hipel, and Kilgour (Fang et al., 1993). In this chapter, the Graph Model for Conflict Resolution and its associated software GMCR I and its descendent GMCR II are reviewed.

2.1 Graph Model for Conflict Resolution: Structure and Implementation

The general steps to follow when applying the graph model for conflict resolution to a real life dispute are shown in Figure 2.1. As depicted in this figure, there are two stages to studying a conflict: modeling and analysis. In the modeling stage, the key decision makers (DMs), options for each DM, and each DM's preferences are identified. In the analysis stage, stability calculations can be carried out manually (for relatively small conflicts) or using the GMCR I or GMCR II software to determine the stable states for each DM and the equilibria.

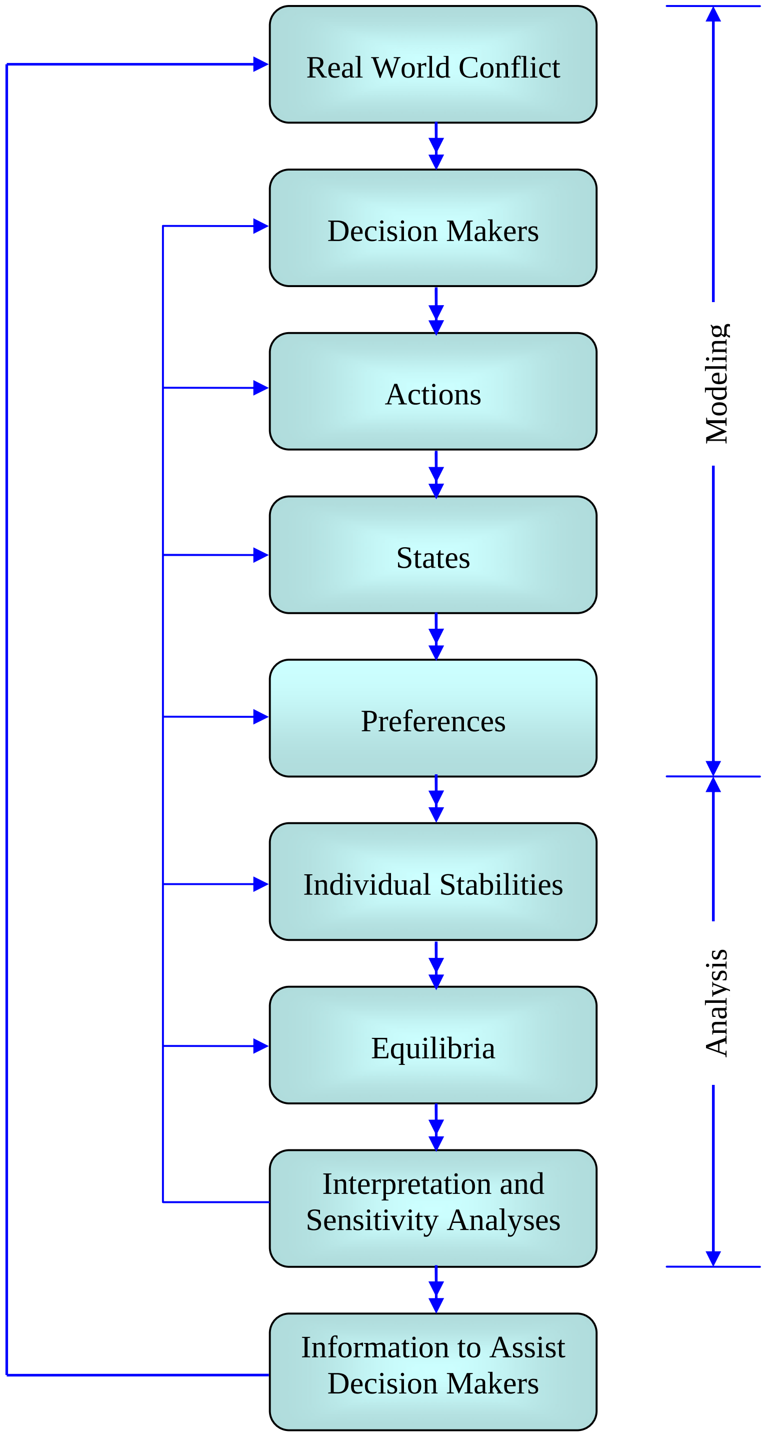

Figure 2.1: General steps for applying the Graph Model for Conflict Resolution.

Fang et al. (1993), represent the states as nodes of a graph and the possible DM's moves as arcs. Assuming that the set of DMs is represented by $N = \{1, 2, ..., n\}$ and the set of states is denoted by $S = \{s_1, s_2, ..., s_q\}$, then the graph model can represented as the collection of the finite directed graphs $\{D_i = (S, A_i), i \in N\}$ where D_i is DM *i*'s directed graph and A_i is the set of directed arcs of DM *i*.

On the set of feasible states represented by the Cartesian product $(S \times S)$, each DM can express his or her relative preferences. DM *i* can prefer state s_1 to s_2 or be indifferent between the two using the notation $s_1 \succ_i s_2$ and $s_1 \sim_i s_2$, respectively.

Each DM controls the moves in his or her own directed graph. The direction of the move from one state to another is shown using directional arrows. The common vertices within all the DMs' graphs constitute the feasible states. From this graph, one can identify the reachable list of states by a specific DM. For DM *i*, the reachable list can be defined as:

$$R_i = \{s_k \in S : (s, s_k) \in A_i\} \qquad 2.1$$

In this case, the reachable list for DM *i*, is the set of states reachable in one step from state s by DM *i*. If $s_k \succ_i s$, then the move by DM *i* constitutes a Unilateral Improvement (UI). Hence, the set of UIs by DM *i* can be defined as:

$$R_i^+ = \{s_k \in R_i(s) : s_k \succ_i s\} \qquad 2.2$$

If $s \succ_i s_k$ or $s \sim_i s_k$, then the set of moves by DM *i* constitutes a set of states that are less or equally preferred to state *s*. Therefore, this set of states can be defined as:

$$R_i^- = \{s_k \in R_i(s) : s \geq_i s_k\} \qquad 2.3$$

Some of the key characteristics of Graph Model for Conflict Resolution are:

- Can handle any finite number of DMs and states.
- Can handle reversible and irreversible moves.
- Can handle common moves for which two or more of the DMs can unilaterally cause the conflict to move from an initial state to the same final state.
- Is easy to visualize and inspect.
- Can support a variety of solution concepts.

The example of a potential nuclear power confrontation shown in Figure 2.2 (Fang et al., 1993) is used to better explain these advantages. In this example, each of the two DMs has three options: peace (P), conventional attack (C), and nuclear war (W for nuclear winter). The first entry on each node represents the strategy choice by DM 1 while the second entry stands for the strategy choice by DM 2. The graph displays some of the reversible moves like, the move between PP and CP in DM 1's graph. It also shows the concepts of irreversible and common moves as depicted by all the arrows going to the node W on both graphs.

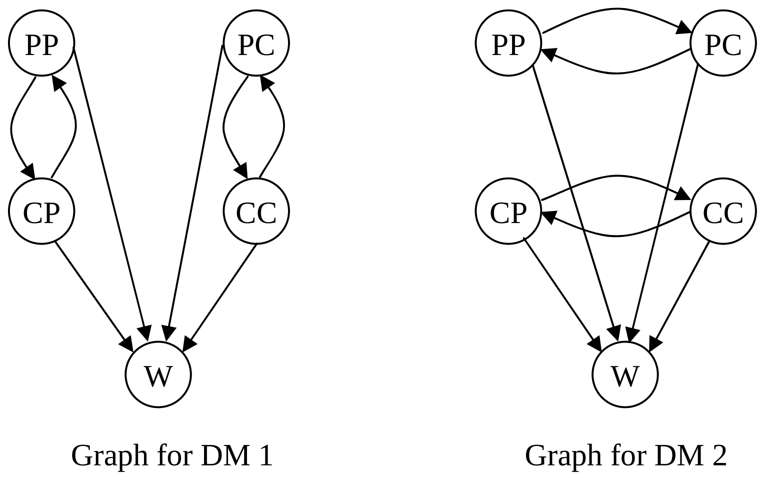

Figure 2.2: Nuclear power confrontation.

2.2 Stability Analysis: Concepts and Definitions

A state is said to be stable with respect to DM *i* if and only if DM *i* is not willing to deviate from it. Whenever the same state is stable for all DMs, it is said to be an equilibrium and may be a possible resolution for the conflict under study. The solution concepts of Nash, General Metarationality (GMR), Symmetric Metarationality (SMR), and Sequential stability (SEQ) are now defined for the case of two DMs:

Definition 2.3.1 Nash stability (Nash): *For* $i, j \in N$, *a state* $s \in S$ *is Nash stable for DM i, denoted by* $s_i^{Nash} \in S$, *iff* $R_i^+(s) = \phi$.

Under Nash stability, DM *i* will always make use of any possible UIs without taking the opponent's move into consideration. Hence, a state *s* is Nash stable for DM *i* iff *i* has no further UIs from state *s*.

Definition 2.3.2 General metarational (GMR): *For* $i, j \in N$, *a state* $s \in S$ *is general metarationally stable for DM i, denoted by* $s_i^{GMR} \in S$, *iff for every* $s_1 \in R_i^+(s)$ *there exists* $s_2 \in R_j(s_1)$ *such that* $s_2 \in R_i^-(s)$.

Hence, every potential UI by DM *i* can be sanctioned by a unilateral move by DM *j* regardless of DM *j*'s preferences. For the situation in which DM *i* has no UI from state s, the state is Nash stable which also implies GMR stability.

Definition 2.3.3 Symmetric metarational (SMR): *For* $i, j \in N$, *a state* $s \in S$ *is symmetric metarationally stable for DM i, denoted by* $s \in S_i^{SMR}$, *iff for every* $s_1 \in R_i^+(s)$ *there exists* $s_2 \in R_j(s_1)$ *such that* $s_2 \in R_i^-$ *and* $s_3 \in R_i^-(s)$ *for all* $s_3 \in R_i(s_2)$.

Accordingly, every potential UI by DM *i* can be sanctioned by a unilateral move levied by DM *j* regardless of DM *j*'s preference, and DM *i* cannot escape from this sanction by a countermove. For the situation in which DM *i* has no UI from state s, the state is Nash stable which also implies SMR stability.

Definition 2.3.4 Sequential stability (SEQ): *For* $i, j \in N$, *a state* $s \in S$ *is sequentially stable for DM i, denoted by* $s \in S_i^{SEQ}$, *iff for every* $s_1 \in R_i^+(s)$ *there exists* $s_2 \in R_j^+(s_1)$ *such that* $s_2 \in R_i^-(s)$.

Hence, every potential UI by DM *i* can be sanctioned by a UI by DM *j*. For the situation in which DM *i* has no UI from state *s*, the state is Nash stable which also implies SEQ stability.

In the two DMs case one needs to consider only the response of one opponent while in the n-DMs case one needs to consider a move or group of moves by all opponents. The solution concepts of Nash, General Metarationality (GMR), Symmetric

Metarationality (SMR), and Sequential stability (SEQ) are now defined for the case of n-DMs:

Definition 2.3.5 Nash stability (Nash): *For* $i \in N$, *a state* $s \in S$ *is Nash stable for DM i, denoted by* $s_i^{Nash} \in S$, *iff* $R_i^+(s) = \phi$.

Under Nash stability, DM *i* will always make use of any possible UIs without taking the opponent's move into consideration. Hence, a state *s* is Nash stable for DM *i* iff *i* has no further UIs from state *s*.

Definition 2.3.6 General metarational (GMR): *For* $i \in N$, *a state* $s \in S$ *is general metarationally stable for DM i, denoted by* $s_i^{GMR} \in S$, *iff for every* $s_1 \in R_i^+(s)$ *there exists at least one* $s_2 \in R_{N-i}(s_1)$ *such that* $s_2 \in R_i^-(s)$.

Hence, every potential UI by DM *i* can be sanctioned by a move by the opponents $N-i$ regardless of $N-i$'s preferences. For the situation in which DM *i* has no UI from state s, the state is Nash stable which also implies GMR stability.

Definition 2.3.7 Symmetric metarational (SMR): *For* $i \in N$, *a state* $s \in S$ *is symmetric metarationally stable for DM i, denoted by* $s \in S_i^{SMR}$, *iff for every* $s_1 \in R_i^+(s)$ *there exists at least one* $s_2 \in R_{N-i}(s_1)$ *such that* $s_2 \in R_i^-$ *and* $s_3 \in R_i^-(s)$ *for all* $s_3 \in R_i(s_2)$.

Accordingly, every potential UI by DM *i* can be sanctioned by a move levied by $N-i$ regardless of $N-i$'s preferences, and DM *i* cannot escape from this sanction by a countermove. For the situation in which DM *i* has no UI from state s, the state is Nash stable which also implies SMR stability.

Definition 2.3.8 Sequential stability (SEQ): *For* $i \in N$, *a state* $s \in S$ *is sequentially stable for DM i, denoted by* $s \in S_i^{SEQ}$, *iff for every* $s_1 \in R_i^+(s)$ *there exists* $s_2 \in R_{N-i}^+(s_1)$ *such that* $s_2 \in R_i^-(s)$.

Hence, every potential UI by DM *i* can be sanctioned by a UI by DMs *N-i*. For the situation in which DM *i* has no UI from state *s*, the state is Nash stable which also implies SEQ stability.

Table 2.1 lists all the above solution concepts along with their behavioral patterns like, foresight, disimprovements, knowledge of preferences, and the risk attitudes.

Table 2.1: Solution concepts and their behavioral patterns.

Solution Concepts	Stability Descriptions	Characteristics			
		Foresight	Disimprovement	Knowledge of Preferences	Strategic Risk
Nash Stability	A focal DM cannot unilaterally move to a more preferred state.	Low	Never	Own	Ignores risk
General Metrationality	All of the focal DM's unilateral improvements are sanctioned by subsequent unilateral moves by others.	Medium	By Opponent	Own	Avoid risk; conservative
Symmetric Metarationality	All focal DM's unilateral improvements are still sanctioned even after possible responses by the focal DM.	Medium	By Opponent	Own	Avoid risk; conservative
Sequential Stability	All of the focal DM's unilateral improvements are sanctioned by subsequent unilateral improvements by others.	Medium	Never	All	Takes some risks; strategizes.

2.3 Stability Analysis using GMCR I and II Software

The graph model for conflict resolution software known as GMCR I was first introduced in the year 1993 as a DOS-based solver (Fang et al., 1993). It provides stability analysis under different types of solution concepts. The second generation of the software, known also as GMCR II, was introduced in 1997 (Hipel, et al., 1997; Fang et al., 2003a, b; Peng, 1999). It was programmed in the C language and works under a Windows environment. GMCR II is more user-friendly and possesses more interactive menus than GMCR I. It provides a variety of information such as individual stability, overall equilibria, and possible coalitions.

When using GMCR II, the user is asked to input the DM, their options, specify the preferences, and the possible combination of infeasibilities. The system will then remove the infeasible states automatically, generate the ordinal preferences, the individual stabilities, and the overall equilibria under different solution concepts. The flowchart of the GMCR II system is shown in Figure 2.3. It has been designed to allow the user to easily setup, update, and modify the conflict under study. GMCR II allows the user to retrieve a previously modeled conflict and give the analyst the ability to change the model parameters to analyze the overall effects on the conflict equilibria.

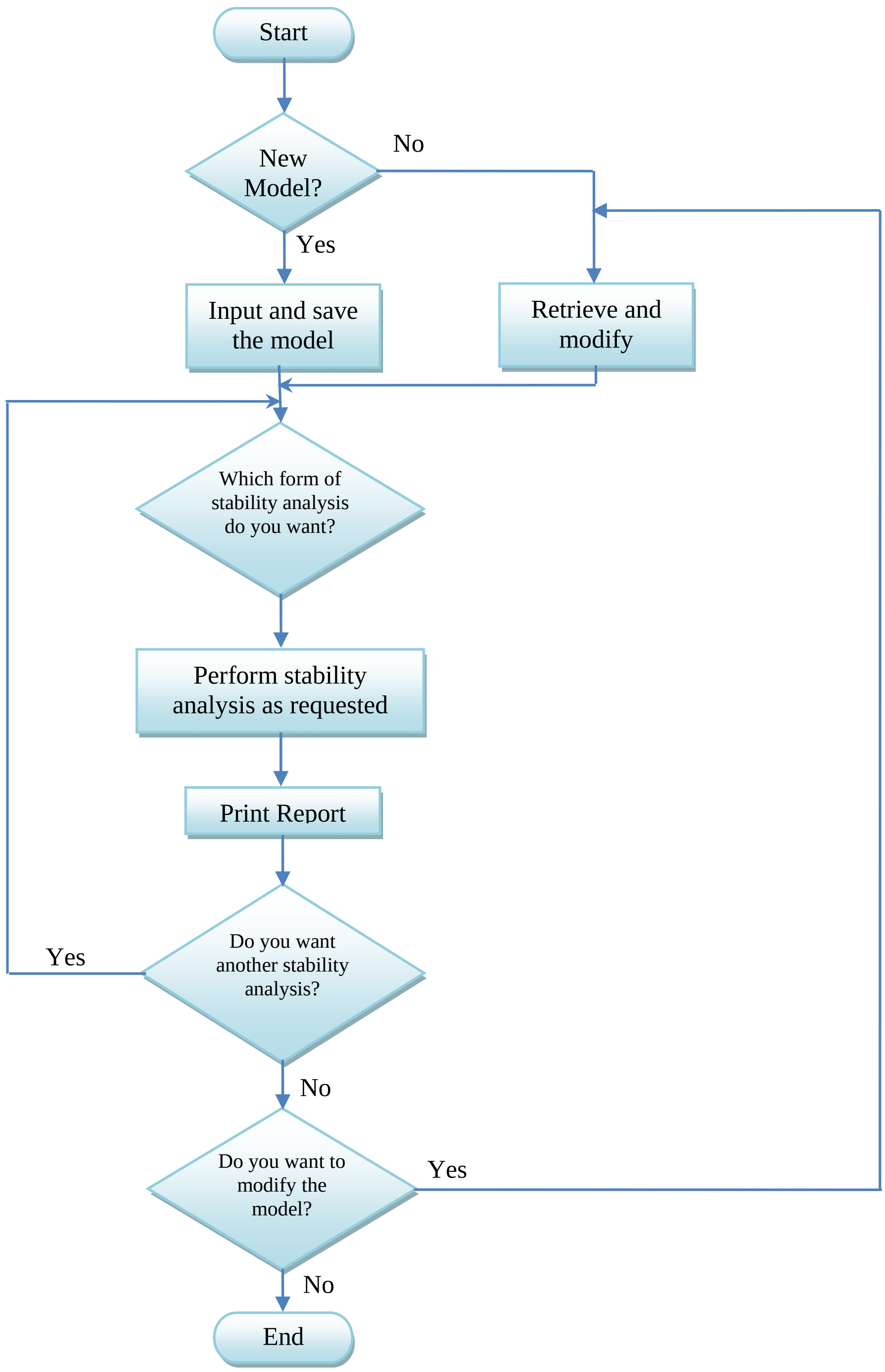

Figure 2.3: Flowchart of GMCR II system.

2.4 Summary

An overview of the graph model for conflict resolution and its associated software GMCR I and GMCR II are covered in this chapter. A flowchart of the GMCR II system is presented in Figure 2.3. Definitions of the solution concepts of Nash, GMR, SMR, and SEQ are also given for the case of two DMs. In Chapters 3 and 4, these solution concepts are expanded to accommodate preference uncertainties.

Chapter 3

Uncertain Preferences in Conflicts with two Participants

Expressing the preferences of the DMs involved in a conflict or a dispute over the set of states is important for predicting a realistic resolution for that conflict. However, in practice, it is not always the case that these preferences can be easily identified. There are cases where the preferences may be uncertain or vague. This vagueness and impreciseness in preferences could influence the predicted resolution or equilibria.

This chapter introduces new definitions for the known solution concepts of Nash, GMR, SMR, and SEQ using fuzzy logic in order to account for the preferences' uncertainties for conflicts with two decision makers. Interrelationships among the newly defined solution concepts are also investigated. The well-known game of Prisoner's Dilemma is employed to illustrate how those new solution concepts can be applied in practice.

3.1 Preference Structures

Determining the stability of a specific state for a given decision maker (DM) depends mainly on the preferences of the DM. Preferences are often expressed using cardinal values or ordinal preferences. Using cardinal values (payoffs) reveals more information than ordinal preferences. While ordinal or relative preferences only express the DM's ordering of alternatives or states, cardinal payoffs reveal the ordering and the strength of the preference among the states. Since ordinal preferences are implied by cardinal values, any model that can handle ordinal preferences can also accommodate cardinal values. The graph model for conflict resolution (Fang et al., 1993) requires only relative preference information.

Pairwise preference information constitutes the most basic form of preference information. For the case of the graph model for conflict resolution, the preference assumptions are as follows. Given $s_1, s_2 \in S$, where S is the set of states, decision maker i (DM i) either strictly prefers s_1 to s_2 (denoted as $s_1 >_i s_2$) or is indifferent between the two (denoted as $s_1 \sim_i s_2$). Those binary relationships can be characterized by the following properties:

1. $>$ is asymmetric: for all $s_1, s_2 \in S$, $s_1 >_i s_2$, and $s_2 >_i s_1$ cannot hold simultaneously.

2. ~ is reflexive: for any $s_1 \in S$, $s_1 \sim_i s_1$.

3. ~ is symmetric: for all $s_1, s_2 \in S$, if $s_1 \sim_i s_2$ then $s_2 \sim_i s_1$.

4. {$>$, ~} is complete: for all $s_1, s_2 \in S$, exactly one of $s_1 >_i s_2$, $s_2 >_i s_2$, or $s_1 \sim_i s_2$ is true.

In the ordinal preference definition, the states are ordered or ranked from most to least preferred, where ties are allowed. When the preferences of some states are the same or when the preferences are not known for sure, the states are said to be equally preferred. In other words, the DM is not in favor of one state over the other. When the states are ranked from most to least preferred with no equally preferred states, the ranking of states is said to be strictly ordinal.

Transitivity is a common underlying preference assumption. If the DM prefers state p to state q, and state q to r, this implies that p is more preferred to r. The ordinal preference definition assumes transitivity.

In cardinal preferences, the payoffs of each state are expressed using real numbers. Often cardinal utility theory (Von Neumann and Morgenstern, 1944) is assumed whereby each state is assigned a utility value. Similar to ordinal preferences, transitivity is one of the underlying assumptions of cardinal preferences.

Generally speaking, imprecision is a key issue when it comes to human judgment, in general, and expressing preferences, in particular. In some situations, a DM's preferences are unclear or imprecise with respect to two or more states. This imprecision in preferences may affect the overall equilibrium predicted for a specific conflict. Acknowledging this lack of precision and trying to handle it using a fuzzy logic approach, is one key contribution of this research.

3.2 Fuzzy Preference Structure

When DM's preferences on a set of states are vague or imprecise, fuzzy approaches provide a flexible framework for handling them (Li et al., 2001; Tanino, 1988; Tanino, 1990). The two main streams in preferences elicitation are binary and non-binary representations. In a binary representation, the preferences are compared at the most basic level. Each of two states or alternatives, as described in the previous section, are compared in a pairwise fashion. The non-binary representation could include both relative preferences (ordinal) and cardinal values (payoffs). In the non-binary case, the overall preferences among the different states are taken into consideration rather than just a pairwise one (Barrett et al., 1990).

3.2.1 Foundational Definitions

Let $S = \{s_1, s_2, \ldots, s_n\}$be a given set of states. Assuming that R is a matrix of binary fuzzy relationships on the Cartesian product $S \times S$, then $r(s_i,s_j)$ or, written simply as r_{ij} denotes the relative degree of preference of state *i* over state *j*. Following this idea, the information matrix, R, is constructed using a square matrix with *i* rows and *j* columns where the diagonal ($i = j$) is always an indifferent preference. A fuzzy function is needed to execute the mapping which associates a fuzzy membership value for each value in the input space (usually a real line) for each state.

The fuzzy function is characterized by a function F to do the mapping into [0,1] where $F(x,\mu_x)$ denotes the relative degree of preference μ_x for state *s*, and *x* may represent a cardinal preference value for a specific state and DM. In this research, a simple triangular form of the function F is assumed as shown in Figure 3.1.

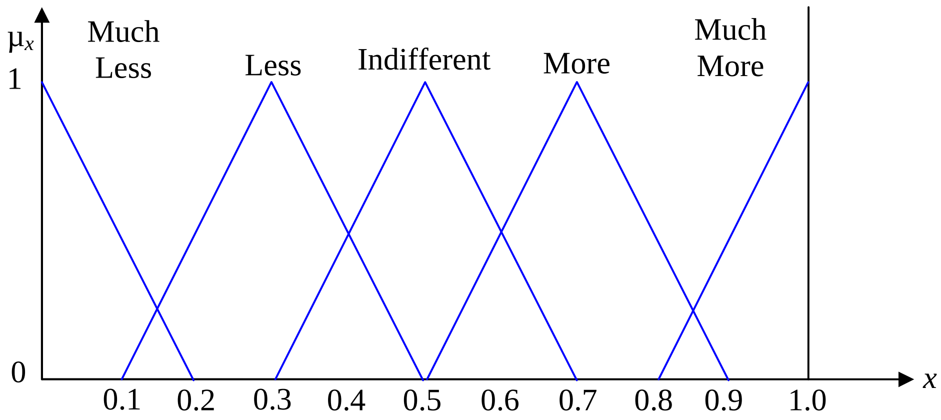

Figure 3.1: Preference representation using five triangular fuzzy sets.

The fuzzy domain of preferences can be divided into five regions with the following linguistic labels:

Much More (MM): refers to all the states that are strongly preferred.

More (M): means all the states that are more preferred.

Indifferent (I): refers to all the states that are equally preferred.

Less (L): means all the states that are less preferred.

Much Less (ML): refers to all the states that are strongly less preferred.

Using the new fuzzy preference structure, for the state $s \in S$, a particular DM *i*, can identify the following subsets with respect to *s*:

$\Phi_i^{MM}(s) = \{s_m : s_m \gg_i s\}$ is the set of states that are much more preferred by DM *i* to state *s*. In this case, s_m and *s* belong to two different fuzzy sets. For example, the values 0.3 and 0.7 on the horizontal axis in Figure 3.1 belong to the fuzzy sets “less” and “more” preferred, respectively.

$\Phi_i^{M}(s) = \{s_m : s_m >_i s\}$ is the set of states that are more preferred by DM *i* to state *s*. The states s_m and *s* belong to the same fuzzy set but with different memberships. For example, the values 0.7 and 0.8 on the horizontal axis in Figure 3.1 both belong to the fuzzy set "more" but with different memberships of 1.0 and 0.5, respectively, as indicated on the vertical axis.

$\Phi_i^I(s) = \{s_m : s_m \sim_i s\}$ is the set of states that are equally preferred by DM i to state s. The states s_m and s belong to the same fuzzy set with the same or extremely close memberships.

$\Phi_i^L(s) = \{s_m : s >_i s_m\}$ is the set of states that are less preferred by DM i to state s. The states s_m and s belong to the same fuzzy set but with different memberships.

$\Phi_i^{ML}(s) = \{s_m : s \gg_i s_m\}$ is the set of states that are much less preferred by DM i to state s. The states s_m and s belong to two different fuzzy sets.

All of the states reachable in a single step by a specific DM i from state s are elements of the reachable list $R_i(s)$. The reachable list can be divided into the following subsets:

$$R_i^{MM}(s) = R_i(s) \cap \Phi_i^{MM}(s)$$

$$R_i^{M}(s) = R_i(s) \cap \Phi_i^{M}(s)$$

$$R_i^{I}(s) = R_i(s) \cap \Phi_i^{I}(s)$$

$$R_i^{L}(s) = R_i(s) \cap \Phi_i^{L}(s)$$

$$R_i^{ML}(s) = R_i(s) \cap \Phi_i^{ML}(s)$$

The set of Unilateral Improvements (UIs) for a specific DM i from a particular state s are:

$$R_i^{M}(s) \cup R_i^{MM}(s).$$

The membership functions of the five fuzzy sets in Figure 3.1 are given by:

$$\mu_{ML} = \begin{cases} -5x + 1 & 0 \le x \le 0.2 \\ 0 & otherwise \end{cases} \quad 3.1$$

$$\mu_{L} = \begin{cases} 5x - 0.5 & 0.1 \le x \le 0.3 \\ -5x + 2.5 & 0.3 \le x \le 0.5 \\ 0 & otherwise \end{cases} \quad 3.2$$

$$\mu_{I} = \begin{cases} 5x - 1.5 & 0.3 \le x \le 0.5 \\ -5x + 3.5 & 0.5 \le x \le 0.7 \\ 0 & otherwise \end{cases} \quad 3.3$$

$$\mu_{M} = \begin{cases} 5x - 2.5 & 0.5 \le x \le 0.7 \\ -5x + 4.5 & 0.7 \le x \le 0.9 \\ 0 & otherwise \end{cases} \quad 3.4$$

$$\mu_{MM} = \begin{cases} 5x - 4 & 0.8 \le x \le 1.0 \\ 0 & otherwise \end{cases} \quad 3.5$$

3.3 Stability Definitions under Fuzzy Preferences

Under a particular solution concept or stability definition, a state *s* is said to be stable for DM *i* if and only if (iff) DM *i* has no incentive to deviate from state *s*. In other words, DM *i* has no UIs from that state or the DM may end up in a worse situation due to sanctioning by others if he or she takes advantage of any UI. If a state is stable for all DMs, it constitutes an equilibrium state and, therefore, a possible resolution to the conflict. When the incentive for a DM not to deviate from a given state is connected to greatly less preferred states, the stability is said to be strong for that DM – otherwise it is weak. If a particular state is strongly stable for all DMs, the state constitutes a strong equilibrium – otherwise it is weak. This idea is different but analogous to the concept of strong and weak stability proposed by Hamouda et al. (Hamouda et al., 2004). The stability definitions given next, are for a conflict consisting of two DMs *i* and *j* who are members of the set of DMs given by *N*.

Definition 3.3.1 Strong Nash stability (SNash): *For DMs* $i, j \in N$, *a state* $s \in S$ *is strongly Nash stable for DM i, denoted by* $s \in S_i^{SNash}$, *iff* $R_i^M(s) \cup R_i^{MM}(s) = \phi$ *and* $s \in \Phi_i^{MM}(s_k)$ *for all* s_k *where* s_k *is any state k from which DM i can unilaterally reach state s in one step.*

A state *s* is said to be strongly Nash stable for a particular DM *i* iff *i* has no UIs from state *s* and *s* is much more preferred to all other states reachable from state *s* by DM *i*. In other words, the preference for state *s* doesn't belong to the same fuzzy set or an adjacent overlapping set representing the preference for the remaining reachable states. Since Nash stability does not take into consideration countermoves by any other DM, this definition holds for $|N| \geq 2$.

Definition 3.3.2 Weak Nash stability (WNash): *For DMs* $i, j \in N$, *a state* $s \in S$ *is weakly Nash stable for DM i, denoted by* $s \in S_i^{WNash}$, *iff* $R_i^M(s) \cup R_i^{MM}(s) = \phi$ *and for all* s_k *for which* $s \in \Phi_i^{MM}(s_k) \cup \Phi_i^{MM}(s_k) \cup \Phi_i^I(s_k)$ *there exists at least one* s_k *such that* $s \in \Phi_i^M(s_k) \cup \Phi_i^I(s_k)$ *where* s_k *is any state k from which DM i can unilaterally reach state s in one step.*

A state *s* is said to be weakly Nash stable for a particular DM *i* iff *i* has no UIs from a state *s* and *s* is more preferred to all other states reachable from state *s*. Hence, the preference for state *s* could belong to the same fuzzy set or an adjacent overlapping set representing the fuzzy preference for one of the states reachable from *s*. Since Nash stability doesn't take into consideration the countermoves by any other DM, this definition holds for $|N| \geq 2$.

Definition 3.3.3 Nash stability (Nash): *The set of Nash stable states for DM i is* $S_i^{Nash} = S_i^{SNash} \cup S_i^{WNash}$.

Definition 3.3.4 General metarationality (GMR): *For DMs* $i, j \in N$, *a state* $s \in S$ *is general metarationally stable for DM i, denoted by* $s \in S_i^{GMR}$, *iff for every* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$ *there exists* $s_2 \in R_j(s_1)$ *such that* $s_2 \in \Phi_i^I(s) \cup \Phi_i^L(s) \cup \Phi_i^{ML}(s)$.

Hence, every potential UI by DM *i* can be sanctioned by a unilateral move by DM *j* regardless of DM *j*'s preferences. For the situation in which DM *i* has no UI from state *s*, the state is Nash stable which also implies GMR stability.

Definition 3.3.5 Strong general metarationality (SGMR): *For DMs* $i, j \in N$, *a state* $s \in S$ *is strongly general metarationally stable for DM i, denoted by* $s \in S_i^{SGMR}$, *iff* $s \in S_i^{GMR}$ *and for every* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$ *there exists* $s_2 \in R_j(s_1)$ *such that* $s_2 \in \Phi_i^{ML}(s)$.

Therefore, every potential UI by DM *i* can be strongly sanctioned by a unilateral move levied by DM *j* regardless of DM *j*'s preferences. For the case in which DM *i* has no UI from state *s*, SGMR is defined to exist only if state *s* is strongly Nash stable.

Definition 3.3.6 Weak general metarationality (WGMR): *For DMs* $i, j \in N$, *a state* $s \in S$ *is weakly general metarationally stable for DM i, denoted by* $s \in S_i^{WGMR}$, *iff* $s \in S_i^{GMR}$ *and for at least one* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$, $R_j(s_1) \cap \Phi_i^{ML}(s) = \phi$.

Every potential UI by DM *i* can be sanctioned but there is at least one UI by DM *i* that can only be weakly sanctioned by DM *j*. For the situation in which DM *i* has no UI from state *s*, WGMR is defined to exist only if state *s* is weakly Nash stable.

Definition 3.3.7 Symmetric metarationality (SMR): *For DMs* $i, j \in N$, *a state* $s \in S$ *is symmetric metarationally stable for DM i, denoted by* $s \in S_i^{SMR}$, *iff for every* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$ *there exists* $s_2 \in R_j(s_1)$ *such that* $s_2 \in \Phi_i^I(s) \cup \Phi_i^L(s) \cup \Phi_i^{ML}(s)$ *and* $s_3 \in \Phi_i^I(s) \cup \Phi_i^L(s) \cup \Phi_i^{ML}(s)$ *for all* $s_3 \in R_i(s_2)$.

Accordingly, every potential UI by DM *i* can be sanctioned by a unilateral move levied by DM *j* regardless of DM *j*'s preference, and DM *i* cannot escape from this sanction by a countermove. For the situation in which DM *i* has no UI from state *s*, the state is Nash stable which also implies SMR stability.

Definition 3.3.8 Strong symmetric metarationality (SSMR): *For DMs* $i, j \in N$, *a state* $s \in S$ *is strongly symmetric metarationally stable for DM i, denoted by* $s \in S_i^{SSMR}$, *iff* $s \in S_i^{SMR}$ *and for every* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$ *there exists* $s_2 \in R_j(s_1)$ *such that* $s_2 \in \Phi_i^{ML}(s)$ *and* $s_3 \in \Phi_i^{ML}(s)$ *for all* $s_3 \in R_i(s_2)$.

Every potential UI by DM *i* can be strongly sanctioned by a move levied by DM *j* regardless of DM *j*'s preference, and DM *i* cannot escape from this sanction by a countermove. For the case in which DM *i* has no UI from state *s*, SSMR is defined to exist only if state *s* is strongly Nash stable.

Definition 3.3.9 Weak symmetric metarationality (WSMR): *For DMs* $i, j \in N$, *a state* $s \in S$ *is weakly symmetric metarationally stable for DM i, denoted by* $s \in S_i^{WSMR}$, *iff* $s \in S_i^{SMR}$ *and for at least one* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$, *exactly one of the following conditions is met:*

1. *There exists* $s_2 \in R_j(s_1)$ *such that* $s_2 \in \Phi_i^I(s) \cup \Phi_i^L(s) \cup \Phi_i^{ML}(s)$ *but for at least one* $s_2 \in R_j(s_1)$, $s_2 \notin \Phi_i^{ML}(s)$ *and* $s_3 \in \Phi_i^I(s) \cup \Phi_i^L(s) \cup \Phi_i^{ML}(s)$ *for all* $s_3 \in R_i(s_2)$.
2. *There exists* $s_2 \in R_j(s_1)$ *such that* $s_2 \in \Phi_i^{ML}(s)$ *and* $s_3 \in \Phi_i^I(s) \cup \Phi_i^L(s)$ *for at least one* $s_3 \in R_i(s_2)$.

At least one potential UI by DM *i* will be strongly sanctioned by a move levied by DM *j* but a countermove by DM *i* will weaken the sanction or the sanction is weak in the first place but it is inescapable by DM *i*. For the situation in which DM *i* has no UI from state *s*, WSMR is defined to exist only if state *s* is weakly Nash stable.

Definition 3.3.10 Sequential stability (SEQ): *For DMs* $i, j \in N$, *a state* $s \in S$ *is sequentially stable for DM i, denoted by* $s \in S_i^{SEQ}$, *iff for every* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$ *there exists* $s_2 \in R_j^M(s_1) \cup R_j^{MM}(s_1)$ *such that* $s_2 \in \Phi_i^I(s) \cup \Phi_i^L(s) \cup \Phi_i^{ML}(s)$.

Therefore, every potential UI by DM *i* can be sanctioned by the opponent *j*'s UIs. For the situation in which DM *i* has no UI from state *s*, the state is Nash stable which also implies SEQ stability.

Definition 3.3.11 Strong sequential stability (SSEQ): *For DMs* $i, j \in N$, *a state* $s \in S$ *is strongly sequentially stable for DM i, denoted by* $s \in S_i^{SSEQ}$, *iff* $s \in S_i^{SEQ}$ *and for every* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$ *there exists* $s_2 \in R_j^M(s_1) \cup R_j^{MM}(s_1)$ *such that* $s_2 \in \Phi_i^{ML}(s)$.

Hence, every potential UI by DM *i* can be strongly sanctioned by the opponent *j*'s UIs. For the case in which DM *i* has no UI from state *s*, SSEQ is defined to exist only if state *s* is strongly Nash stable.

Definition 3.3.12 Weak sequential stability (WSEQ): *For DMs* $i, j \in N$, *a state* $s \in S$ *is weakly sequentially stable for DM i, denoted by* $s \in S_i^{WSEQ}$, *iff* $s \in S_i^{SEQ}$ *and for at least one* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$, $(R_j^M(s_1) \cup R_j^{MM}(s_1)) \cap \Phi_i^{ML}(s) = \phi$.

Hence, at least one UI by DM *i* does not produce a much less preferred state as a result of a credible sanction by the opponent DM *j*. For the situation in which DM *i* has no UI from state *s*, WSEQ is defined to exist only if state *s* is weakly Nash stable.

3.4 Interrelationships among Stability Definitions

Knowing the relationships between the different solution concepts can be very useful. Fang et al. (1993) have proven that a state *s* that is Nash is also GMR, SMR, and SEQ. In other words, if *s* satisfies $s \in S_i^{Nash}$ then it also must satisfy $s \in S_i^{GMR}$, $s \in S_i^{SMR}$, and $s \in S_i^{SEQ}$. Likewise, a state *s* that satisfies $s \in S_i^{SNash}$ also satisfies $s \in S_i^{SGMR}$, $s \in S_i^{SSMR}$, and $s \in S_i^{SSEQ}$. On the other hand, a state *s* that satisfies $s \in S_i^{WNash}$ also satisfies $s \in S_i^{WGMR}$, $s \in S_i^{WSMR}$, and $s \in S_i^{WSEQ}$.

By definition, the relationships between the weak and strong solution concepts are as follows:

$S_i^{Nash} = S_i^{WNash} \cup S_i^{SNash}$.

$S_i^{GMR} = S_i^{WGMR} \cup S_i^{SGMR}$.

$S_i^{SMR} = S_i^{WSMR} \cup S_i^{SSMR}$.

$S_i^{SEQ} = S_i^{WSEQ} \cup S_i^{SSEQ}$.

3.5 Fuzzy Binary Preference Relationships

As mentioned earlier, the preferences of a DM can be expressed using pairwise comparisons. In this section, the terminology is modified to incorporate fuzziness in preferences. For the fuzzy relationship to hold on the Cartesian product $S \times S$, the following conditions should be met:

1. Transitivity:

$\exists i, j, k = 1, ..., n, \quad r(s_i, s_j) \geq 0.5, \quad r(s_j, s_k) \geq 0.5 \quad \Rightarrow r(s_i, s_k) \geq \min(r(s_i, s_j), r(s_j, s_k)).$

This is known as the max-min transitivity (Dubois and Prade, 1980).

2. $\forall x_i, x_j \in X$, $0 \leq r(s_i, s_j) \leq 1.$

When $r(s_i, s_j) = 1,$ this indicates that s_i is more preferred to s_j. The interval $0.5 \leq r(s_i, s_j) \leq 1$ denotes the definite preference interval. When $r(s_i, s_j) = 0.5,$ this means that s_i is equally preferred to s_j. On the other hand, when $r(s_i, s_j) = 0,$ this means that s_i is less preferred to s_j. The interval $0 \leq r(s_i, s_j) \leq 0.5$ denotes the definite less preference interval. The membership functions of the three fuzzy sets in Figure 3.2, where LP, I, and MP refer to less preferred, indifferent, and more preferred, respectively, are given as:

$$\mu_{LP} = \begin{cases} -2x + 1 & 0 \le x \le 0.5 \\ 0 & otherwise \end{cases} \quad 3.6$$

$$\mu_{I} = \begin{cases} 4x - 1 & 0.25 \le x \le 0.5 \\ -4x + 3 & 0.5 \le x \le 0.75 \end{cases} \quad 3.7$$

$$\mu_{MP} = \begin{cases} 2x - 1 & 0.5 \le x \le 1 \\ 0 & otherwise \end{cases} \quad 3.8$$

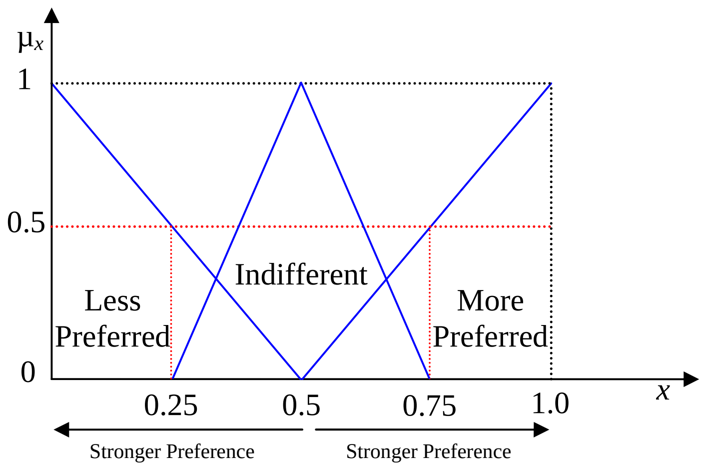

Figure 3.2: Fuzzy binary relationship.

3.6 Fuzzy versus Non-fuzzy Preferences

Non-fuzzy preferences could be considered as having a fuzzy value of one which means zero fuzziness (crisp). Hence, the non-fuzzy case will always be a subset of the fuzzy one (see Figure 3.3). When representing a fuzzy preference using a single value rather than a set, one is actually referring to the crisp value. The preference relations

$r(s_i, s_j) = 0, 0.5,$ and 1 correspond to the strict preferences consisting of less preferred, equally preferred, and more preferred, respectively, for all $i, j = 1, \ldots, n$ $(\forall i, j = 1, \ldots, n)$. In other words, when $r(s_i, s_j) = 1,$ this indicates that s_i is more preferred to s_j and it is denoted as $s_i > s_j$ in a non-fuzzy representation. Any value in the interval $0.5 \le r(s_i, s_j) \le 1$ indicates a definite preference but with different relative degrees. When $r(s_i, s_j) = 0.5,$ this means that s_i is equally preferred to s_j and the notation in a non-fuzzy representation is $s_i \sim s_j$.

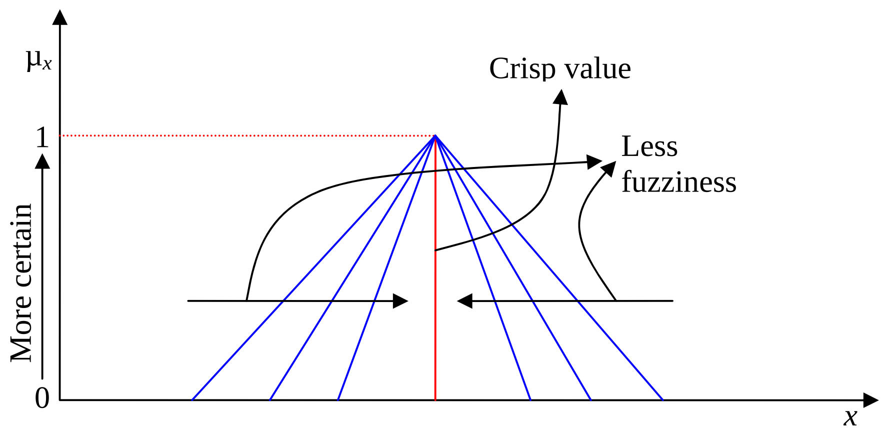

Figure 3.3: Fuzzy non-fuzzy relationship.

When $r(s_i, s_j) = 0,$ this means that s_i is less preferred to s_j and the notation in a non-fuzzy representation is $s_j > s_i$. Any value in the interval $0 \le r(s_i, s_j) \le 0.5$ definitely indicates less preference but with different relative degrees. The fuzzy max-min transitivity implies the usual transitivity: $s_i \ge s_j$, $s_j \ge s_k$, $\Rightarrow s_i \ge s_k$, $\forall i, j, k \in \{1, 2, \ldots, n\}$. For the case of having mixed fuzzy and non-fuzzy preferences, non-fuzzy ones are fuzzified with a membership of one.

3.7 Prisoner's Dilemma from a Fuzzy Perspective

In 1984, Axelrod (Axelrod, 1984) analyzed cooperation by means of a 2×2 non-zero-sum game called "Prisoner's Dilemma". In this game, the two DMs have two strategies: either “cooperate” (called strategy C) or “defect” (labeled as strategy D). Thus, the game is called a 2×2 game. While both DMs gain equally when cooperating, if only one of them cooperates, the other one who defects, will gain more. If both defect, both lose (or gain very little) (Binmore, 1992). Table 3.1 summarizes the complete game situation and its different states or outcomes. Notice, in this table that DM 1, or prisoner 1, controls the row strategies while DM 2 controls the column strategies. When each DM selects a strategy, a state is formed, which is represented by a cell in the matrix. The double letters given at the top of a cell represent the strategies of the DMs where the letters on the left and right stand for the strategies of DMs 1 and 2, respectively. Hence, the cell given as CD which will be referred to by the encircled number 2 for simplicity sake, is the state in which DM 1 cooperates and DM 2 defects. The two numbers given in brackets in the middle of a cell represent the preferences of DM 1 (left entry) and DM 2 (right entry), where a higher number means more preferred. The hypothetical quantities given in brackets at the bottom of a cell are meant to represent the cardinal preference values, where the left and right entries are for DMs 1 and 2, respectively. The higher the number, the fewer years in prison.

The following hypothetical situation is the reason behind the game's name. Two criminals have been arrested under the suspicion of having committed a crime together. Due to lack of evidence, the police can't convict them. While keeping them separated, the police separately offer each of them a deal. The one helping the police to convict the other one will be set free. If they cooperate with each other (by rejecting the police’s offer), both of them will be jailed for a short time and they both will gain the same amount. However, each has an incentive to confess to the police in order to be set free. In this case, the defector will gain more, since he or she will be freed; the one who remains silent, on the other hand, will receive the full punishment. If both comply with the police’s request, both will be jailed, but for less time than if one had refused to talk and the other had confessed. The dilemma arises because each prisoner needs to make a wise decision which is not possible without knowing the other’s choice.

Table 3.1: Prisoner's Dilemma in normal form.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC (3, 3)[1] (6, 6)[2]	② CD (1, 4) (1,10)
	Defect D	③ DC (4, 1) (10,1)	④ DD (2, 2) (3, 3)

1. Relative preferences (ordinal).
2. Cardinal payoffs.

In many real life situations, one can encounter such distributions of gains and losses. While the cooperator, whose action is not reciprocated, will lose resources to the defector, neither of them will be able to get the additional gain coming from their cooperation. The gain for mutual cooperation in Prisoner's Dilemma is kept smaller than the gain for one-sided defection so that there would always be an incentive to defect, though this assumption might not be generally valid. For example, it is better when two hunters together hunt a stag rather than hunting individually as they would for a hare. Even if one of them hunts a hare and gave it to the other one, the hunter who did nothing would still have less gain compared to the case in which he or she had helped his or her companion to hunt a stag.

One of the problems associated with Prisoner's Dilemma is rationality. If both DMs were purely rational, they would never cooperate. A rational decision-making process means that a particular DM makes the decision which is best for him or her regardless of the other DM's choices. If the other DM decided to defect, then it is rational for the given DM to defect. While not gaining by doing so, the initial DM is still avoiding the maximum stay in prison. On the other hand, if the other DM decided to cooperate, the rational choice for the particular DM is to defect so he or she will gain more. If both are rational, both will decide to defect, and neither of them will gain anything. On the other hand, if both would "irrationally" decide to cooperate, both would spend less time in jail. The problem now is to find the appropriate mix between the two scenarios.

When knowing nothing about any future interactions, Prisoner's Dilemma is a generic way for studying short term decision-making. Assuming the evolution of experience is a cumulative process, long term cooperation can only evolve after short term ones have been selected, thereby adding small improvements upon each other but without blindly making major blunders.

Despite the rationalism and the inherent selfishness of people and organizations (Dawkins, 1988), some sort of cooperative behavior may occur among different individuals. The motives for these cooperative actions could be based on, but not limited to, some type of trust, expectancy of reciprocation, or conditional behaviors such as "you do this and I will do that". To better model such cooperative behavior, researchers examine what is called repeated Prisoner's Dilemma. This will eventually enable them to track the evolution of such behavior (Bendor, 1993; Bendor et al., 1996; Flake, 1998).

In a single encounter of Prisoner's Dilemma, if the first prisoner defects, the second will defect in response to the first one's defection. On the other hand, if the first one cooperates, the second will defect to get a better payoff. In general, the second player always has an incentive to defect. It could also be claimed that the first player is better to defect in the first place to avoid betrayal by the second one. However, if they both decided to defect regardless of the other's choice, they will always end up in a less preferred situation compared to both cooperating.

When having repeated encounters of the Prisoner's Dilemma game, DMs have to decide between the value of the current encounter and future ones. Players valuing the current encounter more are motivated to defect in favor of a short term benefit. On the other hand, players seeking a long lasting relationship would be more motivated to build a good positive reputation and a trusting relationship (Kollock, 1993). In the following sub-sections, some of the well-known strategies for repeated encounters are explained (Axelrod, 1984; Axelrod and Hamilton, 1981; Binmore and Samuelson, 1982; Boerlijst et al., 1997; Boyd , 1989; Boyd and Lorberbaum, 1987; Hofstadter, 1983; Kollock, 1993; Kraines and Kraines, 1988; Mesterton-Gibbons and Dugatkin,, 1992; Molander, 1985; Mukherji, 1986; Nowak and Sigmund, 1993b, 1992; Nowak and May, 1992; Rapaport and Chammah, 1965; Sigmund, 1993; Smith, 1982).

3.7.1 Tit for Tat (TFT)

Tit for Tat (TFT) is a deterministic strategy. In this strategy, one player starts with a cooperative action and then behaves the same as the other in future steps. In other words, one will defect in the current step if the opponent defected in the previous one. Likewise, one will cooperate in the current step if the opponent did so in the previous step. This strategy is the most well known and has been shown to give the best results in most of the cases. This could be interpreted as a kind of reciprocal cooperation, meaning that one will cooperate based on the expectancy that the other will cooperate too. This is the safest and most rewarding in repeated encounter situations (Axelrod and W. Hamilton, 1981; Hofstadter, 1983; Mesterton-Gibbons and Dugatkin, 1992).

3.7.2 Generous Tit for Tat (GTFT)

As implied by its name, this strategy is a slightly modified and more generous version of the original deterministic TFT strategy. It is mainly based on the assumption of the presence of some noise or misperception. Mistakes in the application of some choices and misperceptions are two common characteristics of human behavior. When implementing this strategy, a player is more forgiving in case of small mistakes or defections in the absence of strong evidence that it was intentional. Adding some generosity will contribute to more cooperative actions (Molander, 1985; Nowak and Sigmund, 1992).

3.7.3 Contrite Tit for Tat (CTFT)

This is another modified version of the famous deterministic TFT strategy. It has three main attitudes namely, contrite, content and provoked. The initiating player begins with cooperation and keeps cooperating unless there is a defection from the other player. If a player defects while content, the victim player becomes provoked and defects until cooperation from the other player causes him or her to become content again. If the other player is content and the remaining player defects, this player becomes contrite and should cooperate. When contrite, he or she becomes content only after he or she has successfully cooperated (Boerlijst et al., 1997; Boyd, 1989).

3.7.4 Suspicious Tit for Tat (STFT)

This is also a modified version of the deterministic TFT. It is based on total distrust. To be on the safe side and avoid any chances of betrayal, a player will defect on the first move; otherwise he or she will do the same as the other player last did. If one makes the first move and defects against TFT, the result is continues defection thereafter (Boyd and Lorberbaum, 1987; Kollock, 1993).

3.7.5 Tit for Two Tats (TF2T or TFTT)

In this strategy, a player is trying to stay somehow neutral (not going to any of the two extremes). Based on the expectancy of an unintentional error or misperception, the player will cooperate on the first move and defect after two consecutive defections by the opponent. It is a more tolerant strategy but very exploitable by a strategy which alternately cooperates and defects (Axelrod, 1984; Kollock, 1993).

3.7.6 PAVLOV

Pavlov is a stochastic simple win-stay, lose-shift strategy. In brief, if the player's payoff is below a certain level, he or she will change his or her action. Otherwise, the player keeps repeating the previous one. A Pavlov player tries to divide game results in each step into two groups: success or defeat. If his last result belongs to the success category, he or she plays the same move; otherwise he or she chooses another move. A player will cooperate if and only if both the protagonist and opponent played identically in the last round. Pavlov success is based on two main advantages: it can correct occasional mistakes and exploit unconditional cooperators (Kraines and Kraines, 1988; Nowak and Sigmund, 1993b).

3.7.7 Prudent PAVLOV (P-PAVLOV)

This is a modified version of the well-known PAVLOV strategy. The main distinction lies in the fact that a player will only resume cooperation after two rounds of mutual defection. The key advantage of this strategy is that it will allow one to recover from an opponent's erroneous or unintentional defection or a misperceived defection (Boerlijst, 1997).

3.7.8 REMORSE

The remorse strategy is the complement of the forgiving one. A player practicing remorse switches to cooperation after defecting or being in a “bad standing” or if both players cooperated in the last round. Maintaining a record of the opponent’s “standing” can be of help recovering from an opponent's erroneous defection. One can call this strategy an error-correcting one (Boerlijst, 1997).

3.7.9 Always Cooperate (ALLC)

This strategy is based on blind trust. Regardless of the other player’s behavior, the one implementing ALLC will always cooperate. A player employing such a strategy is exploited by others and vulnerable to defection.

3.7.10 Always Defect (ALLD)

While ALLC is based on blind trust, ALLD is founded on extreme suspicion (trust no one). In this case, one will always defect regardless of the other's choice. It might benefit a player in a single encounter but for sure not in the long run or repeated encounters.

3.7.11 GRIM

GRIM is an unforgiving strategy that starts with cooperation until the opponent defects once, and then defects for the rest of the game. It will cooperate if both players cooperated previously but will revert to ALLD if the other player defects. The biggest disadvantage is that it cannot recover from an erroneous or misperceived defection (Binmore and Samuelson, 1992; Boerlijst, 1997).

3.8 Ordinal Case

Assuming that the cardinal payoffs of the states are not known, each DM must decide whether or not to cooperate based on the relative preference for each state. Since ordinal preferences provide only the ranking or the ordering of the states from the most to the least preferred, the preferences of both DMs are expressed using linguistic terms from Figure 3.1 as shown in Table 3.2 where ML, L, I, M, and MM refer to much less, less, indifferent, more, and much more, respectively.

Table 3.2: Prisoner's Dilemma in normal form using ordinal preferences.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC M,M	② CD L,MM
	Defect D	③ DC MM,L	④ DD I,I

3.8.1 Nash stability

To asses the Nash stability, each DM considers only his or her own preferences. In other words, each DM will compare the preferences of two states at the same time in a binary fashion. DM 1 controls the row strategies and can move in a bidirectional fashion between states 1 and 3 as well as between states 2 and 4. Whether to move or stay should be based on which state is more preferred. For example, from state 1, DM 1 can unilaterally move to state 3. The preference of state 1 is more preferred (M) while the preference of state 3 is much more preferred (MM). To compare the two preferences, one uses equations 3.6, 3.7 and 3.8, and represents each triangular fuzzy set with its middle value (the most probable value, equations 1 through 5). From Figure 3.1, while the middle value for M is 0.7, the middle value for MM is 1. The membership of M to the indifferent and more preferred fuzzy set is 0.2 and 0.4, respectively. Since M belongs to the more preferred fuzzy set with a higher membership, the preference is said to be more preferred rather than indifferent. Likewise, the MM belongs to the more preferred fuzzy set with a membership of 1.0. Notice that both preferences belong to the set of more preferred states but with different memberships. Hence, state 3 is more preferred to state 1 and state 1 is not Nash stable with respect to DM 1. Since both states belong to the same fuzzy set (more preferred), one could conclude that state 3 is weakly stable and denoted as “weak”. Also, from state 2, DM 1 can move to state 4. The preference of state 2 is L while the preference of state 4 is Indifferent (I). L belongs to the less preferred and indifferent fuzzy sets with a memberships of 0.4 and 0.2, respectively. Hence, L is more represented by the less preferred fuzzy set. In a similar way, I belongs to the less preferred, indifferent, and more preferred fuzzy sets with memberships of 0, 1.0, and 0, respectively. Hence, I is represented by the indifferent fuzzy set. Since I belongs to the indifferent set with a membership of 1.0, it is more preferred to the L. Therefore, state 2 is not Nash stable for DM 1. Since both states 1 and 4 belong to two adjacent fuzzy sets that overlap, state 4 is weakly stable. The entire results of the Nash stability analysis for the two DMs are displayed in Table 3.3.

Table 3.3: Prisoner's Dilemma Nash stability using ordinal preferences.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC Unstable,Unstable	② CD Unstable,Weak
	Defect D	③ DC Weak,Unstable	④ DD Weak,Weak

From Table 3.3, it is easy to notice that state 4 constitutes an equilibrium since it is stable for both DMs. Since it is weakly stable for both DMs, one could conclude that it is a weak equilibrium.

3.8.2 General Metarational Stability (GMR)

A state is general metarationally stable for a DM *i* iff for each unilateral improvement (UI), there is a credible sanction by the opponent DM *j*. In other words, if DM *i* tries to improve from the current state, the opponent *j* can move to a new state that is less preferred by DM *i* when compared to the original state regardless of its preference by DM *j*.

For example, from state 1, DM 1 can improve by moving to state 3 (it is a UI as shown in the calculations for Nash stability). From state 3, DM 2 can move to state 4. As per the definition for the general metarational stability, this move is allowed regardless of its payoff for DM 2 as far as it could sanction DM 1. DM 2 will move to state 4 to sanction DM 1. Now the preference of state 4 has to be compared to the preference of state 1 (original state) with respect to DM 1. As shown in Table 3.2, the preferences of states 1 and 4 for DM 1 are M and I, respectively. M belongs to the more preferred set with a membership of 0.4 (see calculations for Nash stability) while I belongs to the indifferent set with a membership of 1.0. Therefore, state 1 is more preferred to state 4 with respect to DM 1. Hence, state 1 is general metarationally stable for DM 1. Since both states 1 and 4 belong to two adjacent overlapping sets, one could conclude that it possesses a weak general metarational stability. The entire general metarational stability analysis is shown in Table 3.4.

Table 3.4: Prisoner's Dilemma general metarational stability using ordinal preferences.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC Weak,Weak	② CD Unstable,Weak
	Defect D	③ DC Weak,Unstable	④ DD Weak,Weak

From Table 3.4, one can see that states 1 and 4 are general metarationally stable for both DMs and hence constitute an equilibrium. Both states constitute a weak general metarational equilibrium and are elements of the set S^{WGMR}.

3.8.3 Symmetric Metarational Stability (SMR)

A state is symmetric metarationally stable for a DM *i* iff for each unilateral improvement (UI), there is a credible sanction by the opponent DM *j* which is inescapable by a countermove by DM *i*. In other words, if DM *i* tries to improve from the current state, the opponent *j* can also move to a new state, regardless of its payoff for DM *j*, that is less preferred by DM *i* when compared to the original state and inescapable by a countermove by DM *i*.

For example, from state 1, DM 1 can improve by moving to state 3 (it is a UI as shown in the calculations for Nash stability). From state 3, DM 2 can move to state 4. Now the preference of state 4 has to be compared to the preference of state 1 (original state) with respect to DM 1. As shown in Table 3.2, the preferences of states 1 and 4 for DM 1 are M and I, respectively. M belongs to the more preferred set with a membership of 0.4 while I belongs to the indifferent set with a membership of 1.0. Therefore, state 1 is more preferred to state 4 with respect to DM 1. DM 1 can only move from state 4 to state 2 which is not a UI. Hence, state 1 is symmetric metarationally stable for DM 1. Since both states 1 and 4 belong to two adjacent overlapping fuzzy sets (indifferent and more preferred), one could conclude that it possesses a weak symmetric metarational stability. The entire sequential stability analysis is shown in Table 3.5.

Table 3.5: Prisoner's Dilemma symmetric metarational stability using ordinal preferences.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC Weak,Weak	② CD Unstable,Weak
	Defect D	③ DC Weak,Unstable	④ DD Weak,Weak

From Table 3.5, one can see that states 1 and 4 are symmetric metarationally stable for both DMs and hence constitute an equilibrium. Both states constitute a weak symmetric metarational equilibrium and are elements of the set S^{WSMR}.

3.8.4 Sequential Stability (SEQ)

A state is sequentially stable for a DM *i* iff for each unilateral improvement (UI), there is a credible sanction by the opponent DM *j*. In other words, if DM *i* tries to improve from the current state, the opponent *j* can also improve to a new state that is less preferred by DM *i* when compared to the original state.

For example, from state 1, DM 1 can improve by moving to state 3 (it is a UI as shown in the calculations for Nash stability). From state 3, DM 2 can move to state 4. As per the definition for the sequential stability, this move is allowed only if it is a UI for DM 2. With respect to DM 2, the preference of state 4 is I while the preference of state 3 is L. In the previous sections, it has been shown that I and L belong to the indifferent and less preferred sets with memberships of 1.0 and 0.4, respectively. Hence, I is more preferred to L and moving to state 4 is considered a UI for DM 2. Now the preference of state 4 has to be compared to the preference of state 1 (original state) with respect to DM 1. As shown in the calculation for the symmetric metarationality, state 1 is more preferred to state 4 with respect to DM 1. Hence, state 1 is sequentially stable for DM 1. Since both states 1 and 4 belong to two adjacent overlapping sets, one could conclude that it possesses a weak sequential stability. The entire sequential stability analysis is shown in Table 3.6.

Table 3.6: Prisoner's Dilemma sequential stability using ordinal preferences.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC Weak,Weak	② CD Unstable,Weak
	Defect D	③ DC Weak,Unstable	④ DD Weak,Weak

From Table 3.6, one can see that states 1 and 4 are sequentially stable for both DMs and hence constitute an equilibrium. Both states constitute a weak sequential equilibrium and are elements of the set S^{WSEQ}.

3.9 Cardinal Case

Assuming that the real payoffs of the states are known, each DM must decide whether or not to cooperate based on the cardinal preference values for each state. The cardinal payoffs for both DMs are shown in Table 3.1 (lower values in each cell). The normalized values are shown in Table 3.7.

Table 3.7: Prisoner's Dilemma in normal form using normalized cardinal values.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC 0.6,0.6	② CD 0.1,1.0
	Defect D	③ DC 1.0,0.1	④ DD 0.3,0.3

3.9.1 Nash Stability

From state 1, DM 1 can move to state 3. The payoff of state 1 is 0.6 while the payoff of state 3 is 1.0. Using equations 6, 7, and 8, while 1.0 belongs to the more preferred set with a membership of 1.0, 0.6 belongs to the indifferent and more preferred sets with memberships of 0.6 and 0.2, respectively. Hence, state 3 is more preferred to state 1 and state 1 is not Nash stable with respect to DM 1. Also, from state 2, DM 1 can move to state 4. The payoff of state 2 is 0.1 while the payoff of state 4 is 0.3. While state 2 has a membership of 0.8 to the less preferred set, state 4 has memberships of 0.4 and 0.2 to the less preferred and indifferent sets, respectively. Both states 2 and 4 belong to the less preferred set but with different memberships, 0.8 and 0.4, respectively, but state 4 is more preferred since it belongs to the set with a smaller membership. Hence, state 2 is not Nash stable for DM 1. The entire results of the Nash stability analysis for the two DMs are shown in Table 3.8.

Table 3.8: Prisoner's Dilemma Nash stability using cardinal values.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC Unstable,Unstable	② CD Unstable,Weak
	Defect D	③ DC Weak,Unstable	④ DD Weak,Weak

From Table 3.8, it is easy to see that state 4 constitutes a weak Nash equilibrium and is therefore an element of S^{WNash} since it is weakly stable for both DMs.

3.9.2 General Metarational Stability (GMR)

From state 1, DM 1 can unilaterally improve by moving to state 3. From state 3, DM 2 can move to state 4. Now the payoff of state 4 has to be compared to state 1 (original state) with respect to DM 1. As shown in Table 3.7, the payoffs of states 1 and 4 for DM 1 are 0.6 and 0.3, respectively. State 1 belongs to the indifferent set with a membership of 0.6 and state 4 belongs to the less preferred set with a membership of 0.4. Hence, state 4 is less preferred when compared to state 1 with respect to DM 1. Since the two states belong to two adjacent overlapping fuzzy sets (less and indifferent), the preference structure in this case is said to be weak preference. Therefore, state 1 is weakly general metarationally stable for DM 1. The entire general metarational stability analysis is shown in Table 3.9.

Table 3.9: Prisoner's Dilemma general metarational stability using cardinal values.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC Weak,Weak	② CD Unstable,Weak
	Defect D	③ DC Weak,Unstable	④ DD Weak,Weak

From Table 3.9, one can see that states 1 and 4 are sequentially stable for both DMs and hence constitute equilibria. Both states 1 and state 4 constitute weak equilibria.

3.9.3 Symmetric Metarational Stability (SMR)

From state 1, DM 1 can unilaterally improve by moving to state 3. From state 3, DM 2 can move to state 4. Now the payoff of state 4 has to be compared to state 1 (original state) with respect to DM 1. As shown in Table 3.7, the payoffs of states 1 and 4 for DM 1 are 0.6 and 0.3, respectively. State 1 belongs to the indifferent set with a membership of 0.6 and state 4 belongs to the less preferred set with a membership of 0.4. Hence, state 4 is less preferred when compared to state 1 with respect to DM 1 and this sanction is not escapable since DM 1 can only move to state 2 which is not a UI for DM 1. Since the two states belong to two adjacent overlapping fuzzy sets (less preferred and indifferent), the preference structure in this case is said to be weak preference. Therefore, state 1 is weakly symmetric metarationally stable for DM 1. The entire sequential stability analysis is shown in Table 3.10.

Table 3.10: Prisoner's Dilemma symmetric metarational stability using cardinal values.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC Weak,Weak	② CD Unstable,Weak
	Defect D	③ DC Weak,Unstable	④ DD Weak,Weak

From Table 3.10, one can see that states 1 and 4 are symmetric metarationally stable for both DMs and hence constitute equilibria. Both states 1 and 4 constitute weak equilibria.

3.9.4 Sequential Stability (SEQ)

From state 1, DM 1 can unilaterally improve by moving to state 3. From state 3, DM 2 can move to state 4 where this move is allowed only if it is a UI for DM 2. With respect to DM 2, the payoff of state 4 is 0.3 while the payoff of state 3 is 0.1. The memberships of 0.3 and 0.1 are $\mu_{LP}(0.3) = 0.4$ and $\mu_{LP}(0.1) = 0.8$, respectively. Hence, state 4 belongs to the less preferred set with a smaller membership and therefore moving to this state is considered a UI for DM 2. Now the payoff of state 4 has to be compared to state 1 (original state) with respect to DM 1. As shown in Table 3.7, the payoffs of states 1 and 4 for DM 1 are 0.6 and 0.3, respectively. State 1 belongs to the indifferent set with a membership of 0.6 and state 4 belongs to the less preferred set with a membership of 0.4. Hence, state 4 is less preferred when compared to state 1 with respect to DM 1. Since the two states belong to two adjacent overlapping fuzzy sets (less preferred and indifferent), the preference structure in this case is said to be weak preference. Therefore, state 1 is weakly sequential stable for DM 1. The entire sequential stability analysis is shown in Table 3.11.

Table 3.11: Prisoner's Dilemma sequential stability using cardinal values.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC Weak,Weak	② CD Unstable,Weak
	Defect D	③ DC Weak,Unstable	④ DD Weak,Weak

From Table 3.11, one can see that states 1 and 4 are sequentially stable for both DMs and hence constitute weak equilibria.

Table 3.12 summarizes the entire stability analysis for Prisoner's Dilemma using the fuzzy approach for both the ordinal and the cardinal cases. In this table, DM 1, and DM 2, and E refer to prisoner 1, prisoner 2, and equilibrium, respectively. The equilibrium states are classified as weak or strong equilibrium. The equilibrium could be strong or weak with respect to all the DMs or just to a particular one.

Table 3.12: Prisoner's Dilemma fuzzy stability analysis.

Stability Type	State	Ordinal DM 1	Ordinal DM 2	Ordinal E	Cardinal DM 1	Cardinal DM 2	Cardinal E
Nash	1	Unstable	Unstable		Unstable	Unstable	
	2	Unstable	Weak		Unstable	Weak	
	3	Weak	Unstable		Weak	Unstable	
	4	Weak	Weak	Weak	Weak	Weak	Weak
GMR	1	Weak	Weak	Weak	Weak	Weak	Weak
	2	Unstable	Weak		Unstable	Weak	
	3	Weak	Unstable		Weak	Unstable	
	4	Weak	Weak	Weak	Weak	Weak	Weak
SMR	1	Weak	Weak	Weak	Weak	Weak	Weak
	2	Unstable	Weak		Unstable	Weak	
	3	Weak	Unstable		Weak	Unstable	
	4	Weak	Weak	Weak	Weak	Weak	Weak
SEQ	1	Weak	Weak	Weak	Weak	Weak	Weak
	2	Unstable	Weak		Unstable	Weak	
	3	Weak	Unstable		Weak	Unstable	
	4	Weak	Weak	Weak	Weak	Weak	Weak

3.10 Summary

A new fuzzy preference structure is proposed to deal with situations in which the preferences of the DM are fuzzy or uncertain. Four solution concepts, Nash, general metarational, symmetric metarational, and sequential stabilities, are redefined to accommodate the fuzziness in preferences for a two-DM conflict. The proposed fuzzy approach is applied to the game of Prisoners' Dilemma. The application demonstrates the effectiveness of the fuzzy approach in modeling conflicts with fuzzy preferences. More specifically, it gives a more realistic way of analyzing the behavior of individuals involved in a conflict, especially when the preferences are vague or imprecise. Using the fuzzy approach allows the preferences to vary within ranges rather than just being single crisp values. Whenever the cardinal preferences are not known, they can be approximated using some linguistic variables. It also allows a DM to consider both the ordinal and cardinal preferences. Overall, the fuzzy approach preference methodology is a more flexible framework for analyzing conflicts with imprecise preferences. In Chapter 4, the solution concepts presented in this chapter are extended to the n-DM cases.

Chapter 4

Uncertain Preferences in Conflicts with Multiple Participants

While in a 2-DM conflicts, one needs to take into account potential responses by the other opponent; in n-DM conflicts, one must consider a move or group of moves by more than one opponent. In this chapter, the new definitions introduced in chapter 3 are extended for use with n-DM conflicts, where $n \geq 2$. Also, the concepts of group unilateral movement (UM) and group unilateral improvement (UI) are introduced. To demonstrate the applicability of the newly proposed definitions, they are applied to an aquifer contamination conflict.

4.1 n-DM Case

The assessment of the stability of an n-DM model requires the examination of all the possible responses by all other DMs to a certain move by a DM $i \in N$ (stability definitions for n-DMs are given in Chapter 2). Defining the concept of an allowable sanction, a key concept in GMR, SMR, and SEQ stability, is the first step in this direction.

Suppose that $H \subseteq N, H \neq \phi,$ is any non-empty subset of DMs, and a state $s \in S$. Starting at a state *s*, $R_H(s) \subseteq S$ refers to the set of states that can result from group unilateral moves (UMs) by *H*. It is important to note that a DM in *H* may move more than once but never two moves successively. If $s_1 \in R_H(s)$, then $\Omega_H(s, s_1)$ will refer to the set of all last DMs in any allowable sequence of moves by *H* from *s* to s_1.

Definition 4.1.1: *Let* $s \in S$ *and* $H \subseteq N$, $H \neq \phi$. *A UM from s by H is a member of* $R_H(s) \subseteq S$, *and is defined inductively by:*

1. *If* $i \in H$ *and* $s_1 \in R_i(s)$, *then* $s_1 \in R_H(s)$ *and* $i \in \Omega_H(s, s_1)$.
2. *If* $s_1 \in R_H(s)$, $i \in H$, *and* $s_2 \in R_i(s_1)$, *then*
    a. *If* $|\Omega_H(s, s_1)| = 1$ *and* $i \notin \Omega_H(s, s_1)$, *then* $s_2 \in R_H(s)$ *and* $i \in \Omega_H(s, s_2)$.
    b. *If* $|\Omega_H(s, s_1)| \succ 1$, *then* $s_2 \in R_H(s)$ *and* $i \in \Omega_H(s, s_2)$.

$R_H(s)$ *and* $R_H^M(s) \cup R_H^{MM}(s)$ *could be thought of as H's UMs and UIs from state s, respectively.*

The reachable list could be attained either by adding states that are single moves from *s* or adding states that are group moves by some or all DMs in the set *H* moving sequentially. The set of all group UIs for *H* from *s* will be denoted by $(R_H^M(s) \cup R_H^{MM}(s)) \subseteq S$ where $R_H^M(s)$ and $R_H^{MM}(s)$ refer to the set of more and much more preferred states by *H* from state *s*.

Definition 4.1.2: *Let* $s \in S$ *and* $H \subseteq N$, $H \neq \phi$. *A UI from s by H is a member of* $(R_H^M(s) \cup R_H^{MM}(s)) \subseteq S$ *and is defined inductively by:*

1. *If* $i \in H$ *and* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$, *then* $s_1 \in R_H^M(s) \cup R_H^{MM}(s)$ *and* $i \in \Omega_H^M(s,s_1) \cup \Omega_H^{MM}(s,s_1)$.

2. *If* $s_1 \in R_H^M(s) \cup R_H^{MM}(s)$, $i \in H$, *and* $s_2 \in R_i^M(s_1) \cup R_i^{MM}(s_1)$, *then*

   a. *If* $\left|\Omega_H^M(s,s_1) \cup \Omega_H^{MM}(s,s_1)\right| = 1$ *and* $i \notin \Omega_H^M(s,s_1) \cup \Omega_H^{MM}(s,s_1)$, *then* $s_2 \in R_H^M(s) \cup R_H^{MM}(s)$ *and* $i \in \Omega_H^M(s,s_2) \cup \Omega_H^{MM}(s,s_2)$.

   b. *If* $\left|\Omega_H^M(s,s_1) \cup \Omega_H^{MM}(s,s_1)\right| > 1$, *then* $s_2 \in R_H^M(s) \cup R_H^{MM}(s)$ *and* $i \in \Omega_H^M(s,s_2) \cup \Omega_H^{MM}(s,s_2)$.

$R_{N-i}(s)$ and $R_{N-i}^M(s) \cup R_{N-i}^{MM}(s)$ represent the group UMs and UIs from state s, respectively by DM *i*'s opponents. $R_H^M(s) \cup R_H^{MM}(s)$ could be attained by either adding states that are single UIs from s or adding states that are group UIs by some or all DMs in the set H.

Definition 4.1.3 Strong Nash stability (SNash): *For DM* $i \in N$, *a state* $s \in S$ *is strongly Nash stable for DM i, denoted by* $s \in S_i^{SNash}$, *iff* $R_i^M(s) \cup R_i^{MM}(s) = \phi$ *and* $s \in \Phi_i^{MM}(s_k)$ *for all* s_k *where* s_k *is any state k from which DM i can unilaterally reach state s in one step.*

A state *s* is said to be strongly Nash stable for a particular DM *i* iff *i* has no UIs from state *s* and *s* is much more preferred to all other states reachable from state *s* by DM *i*. In other words, the preference for state *s* doesn't belong to the same fuzzy set or an adjacent overlapping set representing the preference for the remaining reachable states. Since Nash stability doesn't take into consideration countermoves by any other DM, this definition holds for 2-DMs and any $|N| \geq 2$.

Definition 4.1.4 Weak Nash stability (WNash): *For DM* $i \in N$, *a state* $s \in S$ *is weakly Nash stable for DM i, denoted by* $s \in S_i^{WNash}(s)$, *iff* $R_i^M(s) \cup R_i^{MM}(s) = \phi$ *and for all* s_k *for which* $s \in \Phi_i^{MM}(s_k) \cup \Phi_i^M(s_k) \cup \Phi_i^I(s_k)$ *there exists at least one* s_k *such that* $s \in \Phi_i^M(s_k) \cup \Phi_i^I(s_k)$ *where* s_k *is any state k from which DM i can unilaterally reach state s in one step.*

A state *s* is said to be weakly Nash stable for a particular DM *i* iff *i* has no UIs from a state *s* and *s* is more preferred to all other states reachable from state *s*. Hence, the preference for state *s* could belong to the same fuzzy set or an adjacent overlapping set representing the fuzzy preference for one of the states reachable from *s*. Since Nash stability doesn't take into consideration the countermoves by any other DM, this definition holds for 2-DMs and any $|N| \geq 2$.

Definition 4.1.5 Nash stability (Nash): *The set of Nash stable states for DM i is* $S_i^{Nash} = S_i^{SNash} \cup S_i^{WNash}$.

Definition 4.1.6 General metarationality (GMR): *For DM* $i \in N$, *a state* $s \in S$ *is general metarationally stable for DM i, denoted by* $s \in S_i^{GMR}$, *iff for every* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$ *there exists* $s_k \in R_{N-i}(s_1)$ *such that* $s_k \in \Phi_i^I(s) \cup \Phi_i^L(s) \cup \Phi_i^{ML}(s)$.

Hence, every potential UI by DM *i* can be sanctioned by a move by the opponents $N-i$ regardless of $N-i$'s preferences. For the situation in which DM *i* has no UI from state *s*, the state is Nash stable which also implies GMR stability.

Definition 4.1.7 Strong general metarationality (SGMR): *For DM* $i \in N$, *a state* $s \in S$ *is strongly general metarationally stable for DM i, denoted by* $s \in S_i^{SGMR}$, *iff* $s \in S_i^{GMR}$ *and for every* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$ *there exists* $s_k \in R_{N-i}(s_1)$ *such that* $s_k \in \Phi_i^{ML}(s)$.

Therefore, every potential UI by DM *i* can be strongly sanctioned by a move levied by the opponents $N-i$ regardless $N-i$'s preferences. For the case in which DM *i* has no UI from state *s*, SGMR is defined to exist only if state s is strongly Nash stable.

Definition 4.1.8 Weak general metarationality (WGMR): *For DM* $i \in N$, *a state* $s \in S$ *is weakly general metarationally stable for DM i, denoted by* $s \in S_i^{WGMR}$, *iff* $s \in S_i^{GMR}$ *and for at least one* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$, $R_{N-i}(s_1) \cap \Phi_i^{ML}(s) = \phi$.

Every potential UI by DM *i* can be sanctioned but there is at least one UI by DM *i* that can only be weakly sanctioned by $N-i$. For the situation in which DM *i* has no UI from state *s*, WGMR is defined to exist only if state s is weakly Nash stable.

Definition 4.1.9 Symmetric metarationality (SMR): *For DM* $i \in N$, *a state* $s \in S$ *is symmetric metarationally stable for DM i, denoted by* $s \in S_i^{SMR}$, *iff for every* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$ *there exists* $s_k \in R_{N-i}(s_1)$ *such that* $s_k \cup R_i(s_k) \subseteq \Phi_i^I(s) \cup \Phi_i^L(s) \Phi_i^{ML}(s)$.

Accordingly, every potential UI by DM *i* can be sanctioned by a move levied by $N-i$ regardless of $N-i$'s preferences, and DM *i* cannot escape from this sanction by a countermove. For the situation in which DM *i* has no UI from state *s*, the state is Nash stable which also implies SMR stability.

Definition 4.1.10 Strong symmetric metarationality (SSMR): *For DM* $i \in N$, *a state* $s \in S$ *is strongly symmetric metarationally stable for DM i, denoted by* $s \in S_i^{SSMR}$, *iff* $s \in S_i^{SMR}$ *and for every* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$ *there exists* $s_k \in R_{N-i}(s_1)$ *such that* $s_k \cup R_i(s_k) \subseteq \Phi_i^{ML}(s)$.

Every potential UI by DM *i* can be strongly sanctioned by a move levied by $N - i$ regardless of $N - i$ '*s* preference, and DM *i* cannot escape from this sanction by a countermove. For the case where DM *i* has no UI from state *s*, SSMR is defined to exist only if state *s* is strongly Nash stable.

Definition 4.1.11 Weak symmetric metarationality (WSMR): *For DM* $i \in N$, *a state* $s \in S$ *is weakly symmetric metarationally stable for DM i, denoted by* $s \in S_i^{WSMR}$, *iff* $s \in S_i^{SMR}$ *and for at least one* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$, *the following condition must holds: There exists* $\{s_k\} \subseteq R_{N-i}(s_1)$, *but for at least one* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$, $s_k \cup R_i(s_k) \not\subset \Phi_i^{ML}(s)$ *where* s_k *is any state k from which DM i can unilaterally reach state s in one step.*

At least one potential UI by DM *i* can be strongly sanctioned by a move levied by $N - i$ but a countermove by DM *i* will weaken the sanction or the sanction is weak in the first place but it is inescapable by DM *i*. For the situation in which DM *i* has no UI from state *s*, WSMR is defined to exist only if state *s* is weakly Nash stable.

Definition 4.1.12 Sequential stability (SEQ): *For DM* $i \in N$, *a state* $s \in S$ *is sequentially stable for DM i, denoted by* $s \in S_i^{SEQ}$, *iff for every* $s_1 \in R_i^M(s) \cup R_i^{MM}(s)$ *there exists* $\{s_k\} \subseteq R_{N-i}^M(s_1) \cup R_{N-i}^{MM}(s_1)$ *such that* $s_k \in \Phi_i^I(s) \cup \Phi_i^L(s)\Phi_i^{ML}(s)$.

Every potential UI by DM *i* can be sanctioned by a UI levied by $N-i$. For the situation in which DM *i* has no UI from state *s*, the state is Nash stable which also implies SEQ stability.

Definition 4.1.13 Strong sequential stability (SSEQ): *For DM* $i \in N$, *a state* $s \in S$ *is strongly sequentially stable for DM i, denoted by* $s \in S_i^{SSEQ}$, *iff* $s \in S_i^{SEQ}$ *and for every* $s_1 \in R_i^{M}(s) \cup R_i^{MM}(s)$ *there exists* $\{s_k\} \subseteq R_{N-i}^{M}(s_1) \cup R_{N-i}^{MM}(s_1)$ *such that* $s_k \in \Phi_i^{ML}(s)$.

Hence, every potential UI by DM *i* can be strongly sanctioned by the opponent $N-i$ UIs. For the case in which DM *i* has no UI from state *s*, SSEQ is defined to exist only if state *s* is strongly Nash stable.

Definition 4.1.14 Weak sequential stability (WSEQ): *For DM* $i \in N$, *a state* $s \in S$ *is weakly sequentially stable for DM i, denoted by* $s \in S_i^{WSEQ}$, *iff* $s \in S_i^{SEQ}$ *and for at least one* $s_1 \in R_i^{M}(s) \cup R_i^{MM}(s)$, $\left(R_{N-i}^{M}(s_1) \cup R_{N-i}^{MM}(s_1)\right) \cap \Phi_i^{ML}(s) = \phi$.

Hence, at least one UI by DM *i* does not produce a much less preferred state as a result of a credible sanction by the opponent $N-i$. For the situation in which DM *i* has no UI from state *s*, WSEQ is defined to exist only if state *s* is weakly Nash stable.

4.2 Groundwater Contamination Conflict

In a rich agricultural land in Southern Ontario, Canada, the town of Elmira is situated about 15 kilometers north of the twin cities of Kitchener and Waterloo (see Figure 4.1). It is a prosperous town that is famous for its annual maple syrup festival which is the largest in the world. The main water supplies for the 7400 residences of this small town come mainly from an underground aquifer. In the year 1989, the Ontario Ministry of Environment (MoE), noticed that the water supplies were contaminated with N-nitroso demethylamine (NDMA). Owning a pesticide and rubber plant in the town and known for its bad environmental records, Uniroyal Chemical Ltd (UR) was the main suspect. A Control Order was issued by the MoE requesting UR to implement a long term collection and treatment system. UR cooperation was important in the determination of the cause as well as the best way to cleanse the contaminated aquifer and to carry out the necessary cleaning actions under the supervision of MoE. UR immediately exercised its right to appeal in order to lengthen the process hoping that the Control Order would be canceled or at least modified. The Township of Woolwich and the Regional Municipality of Waterloo (referred to as Local Government or LG for short) were encouraged by their citizens to take a strong position in the dispute. The main decision makers (DMs) involved in this conflict have different objectives which might seem contradicting at some points of time. While the MoE wants to carry out its responsibilities in an effective and efficient way, the UR would like the Control Order to be lifted or modified. On the other hand, the local government wants to protect its citizens and industrial base (Hipel et al., 1993; Kilgour et al., 2001).

Table 4.1 shows the main DMs and their options. The Control Order has already been issued by the MoE but it can still modify it to make it more favorable to UR. For its part, UR can exercise its right to appeal and gain more time, accept the original Control Order as is, or just simply abandon its operations in Elmira. In order to protect its citizens and its industrial base, the LG would insist on the application of the original Control Order.

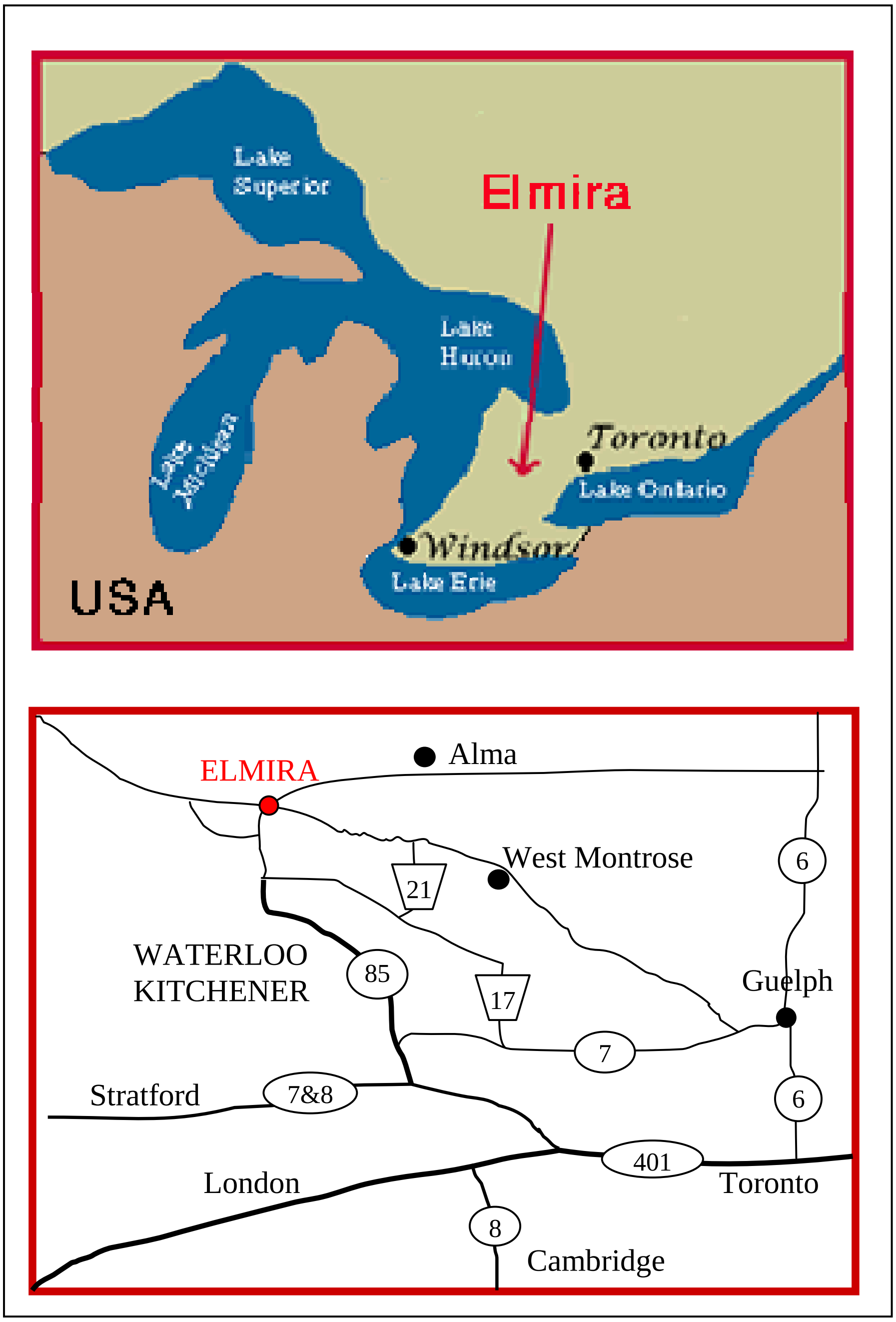

Figure 4.1: Location and map of Elmira.

Table 4.1: Decision makers and their options in the Elmira Conflict.

Decision Maker	Options
Ministry of Environment (MoE)	**1. Modify** the Control Order to make it more acceptable for UR
Uniroyal Chemical Ltd (UR)	**2. Delay** lengthen the appeal process
	3. Accept the original Control Order
	4. Abandon the Elmira operations
Local Government (LG)	**5. Insist** that the original Control Order be applied

Only 9 states are feasible out of 32 possible states. The situation where UR accepts the Control Order or appeals it and at the same time abandons its operations in Elmira is an example of a possible state but not a feasible one. Table 4.2 lists the complete set of feasible states. For simplicity, the feasible states have been numbered 1 through 9. A “Y” besides an option for a particular DM means that the DM chose that option. On the other hand, an “N” besides an option for a particular DM means that the DM decided not to choose that option. When having a “–“ besides an option, it means either a “Y” or a “N”, because it actually doesn’t make a difference whether it is a “Y” or an “N”. To illustrate such a situation, when UR decides to abandon its operations in Elmira, it doesn’t make a difference what other options are selected by any of the remaining DMs.

Table 4.2: Feasible states in the Elmira Conflict.

DM	Options	1	2	3	4	5	6	7	8	9
MoE	1. Modify	N	Y	N	Y	N	Y	N	Y	–
UR	2. Delay	Y	Y	N	N	Y	Y	N	N	–
	3. Accept	N	N	Y	Y	N	N	Y	Y	–
	4. Abandon	N	N	N	N	N	N	N	N	Y
LG	5. Insist	N	N	N	N	Y	Y	Y	Y	–

The preferences of all the DMs over all the possible states are shown in Table 4.3. Assuming that the cardinal or actual payoffs of the states are not known, each DM must express his/her relative preference for each state using some linguistic terms from Figure 1 where ML, L, I, M, and MM refer to much less, less, indifferent, more, and much more, respectively.

Figure 4.2 summarizes the entire Elmira Conflict in what is called an integrated graph model defined within the paradigm of the Graph Model for Conflict Resolution (Fang et al., 1993). It shows all the DMs (labels on the arrows), the feasible states (circled numbers), and the allowable moves by each DM between the different feasible states (directed arrows). For example, state 7 is the much more preferred by both MoE and LG while for UR, state 7 is less preferred. From this state, MoE can only move in one direction towards state 8 (irreversible move represented by a one-sided arrow). Likewise, UR can move from state 7 in an irreversible fashion to state 9. All of the moves in this diagram are irreversible with the exception of the moves by LG as all its moves are reversible.

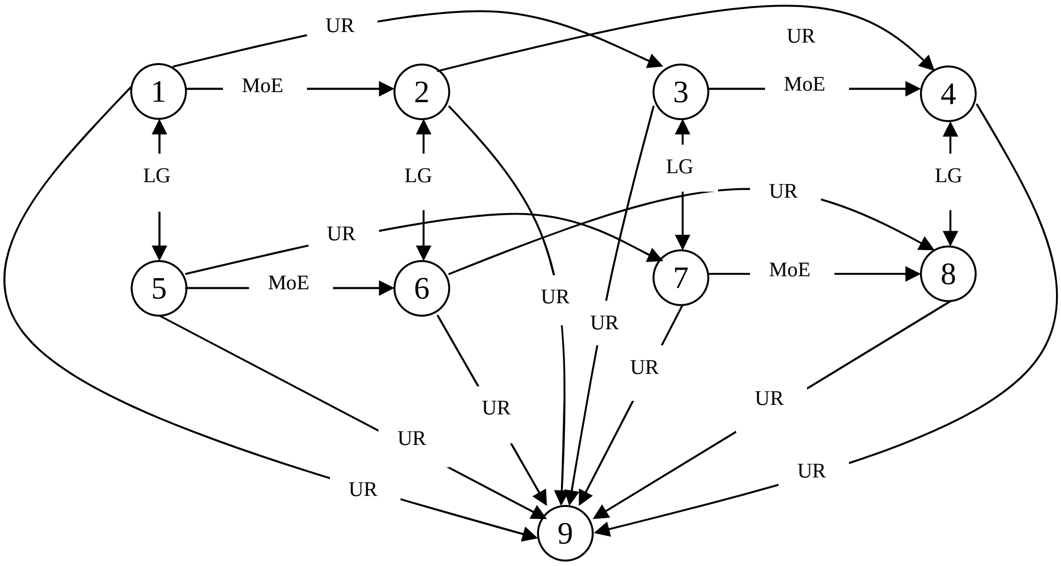

Figure 4.2: Integrated graph of the Elmira Conflict.

Table 4.3: Preferences for each decision maker.

State / DM	1	2	3	4	5	6	7	8	9
MoE	I	L	MM	M	I	L	MM	M	ML
UR	MM	ML	L	MM	M	ML	L	M	I
LG	L	L	MM	I	M	L	MM	I	ML

Doing the analysis manually or using the Graph Model for Conflict Resolution software called GMCR II (see (Fange et al., 2003a, 2003b) for more details), the equilibrium points are states 4, 5, 8, and 9. For example, from state 5, MoE can unilaterally move to state 6. While the preference for state 5 is I, the preference for state 6 is L. I and L belong to the indifferent and less preferred sets with memberships of 1.0 and 0.4, respectively. Hence, state 5 is more preferred to state 6 and moving to state 6 is not a UI for MoE. Therefore, state 5 is Nash stable for MoE. Since the preference for the two states belongs to two adjacent overlapping fuzzy sets, the preference is weak. Hence, state 5 is also a weak GMR, SMR, and SEQ as per the definitions for those solution concepts. Table 4.5 shows the complete stability analysis for the conflict. State 9 is weakly stable for all DMs and hence is the least preferred equilibrium among the four equilibrium points. On the other hand, states 5 and 8 are strongly stable for a single DM namely LG and MoE, respectively. State 4 is the strongest equilibrium as it is strongly stable for both MoE and UR. State 4 represents a possible resolution for the conflict where the MoE modifies the Control Order to make it more favorable for UR which will accept it without any pressure from LG to apply the original one. However, one knows from the conflict literature that LG is pushed by its local citizens to insist on the application of the original control order. Therefore, equilibria at states 5 and 8 are more probable than at state 4.

As depicted in Figure 4.2, it is not possible for any of them to move from state 5 to state 8 on an individual basis. In order to reach state 8, a transition is required either through state 6 or state 7. If MoE moves first to state 6, which is much less preferred by UR, and, subsequently, UR can move from state 6 to state 4. If UR moves first to state 7, which is less preferred by UR but more preferred by both MoE and LG, MoE needs to move to state 8 which constitutes an equilibrium point and hence a possible resolution for the conflict.

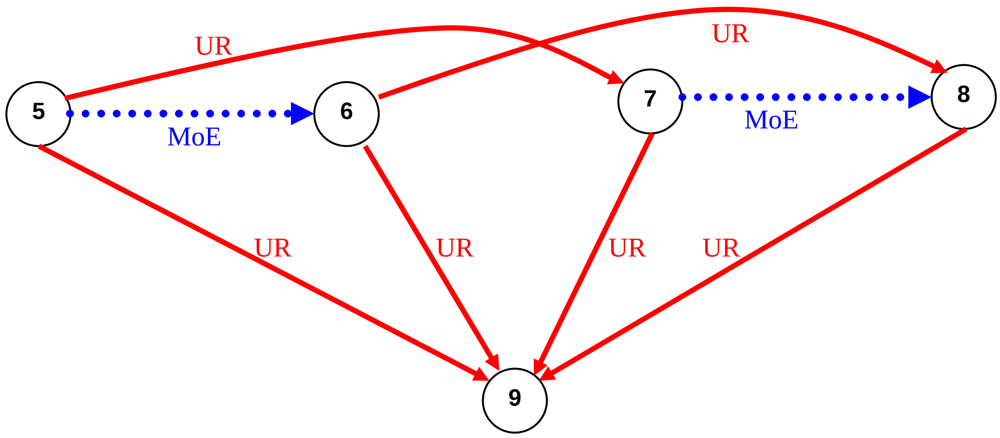

Figure 4.3: State transitions between states 5 and 8.

Table 4.5 shows the evolution of the Elmira conflict from the status quo, to a transitional state and ultimately to a final equilibrium. The status quo is state number 1 which is the most preferred state for UR but less preferred by both MoE and LG. As indicated by the arrow connecting states 5 and 6 in Table 4.5, MoE can unilaterally cause the conflict to move from state 5 to 6 by changing its strategy from insisting that the original Control Order be adopted to modifying it to make it more favorable to UR. From state number 6, UR will move to state 8 which is more preferred by all DMs.

Alternatively, UR could move from state 5 to state 7 as indicated in the bottom of Table 4.5 by changing its strategy to accepting the Control Order. Meanwhile, MoE will modify the original Control Order to make it more favorable to UR and this will cause the entire conflict to end up in state 8.

Table 4.4: Complete stability analysis for the Elmira Conflict.

NASH									
State	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
DM									
MoE	Weak	Weak	Weak	Strong	Weak	Weak	Weak	Strong	Weak
UR	Strong	Unstable	Unstable	Strong	Weak	Unstable	Unstable	Weak	Weak
LG	Unstable	Weak	Weak	Weak	Strong	Weak	Weak	Weak	Weak
EQ				Weak	Weak			Weak	Weak
GMR									
State	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
DM									
MoE	Weak	Weak	Weak	Strong	Weak	Weak	Weak	Strong	Weak
UR	Strong	Unstable	Unstable	Strong	Weak	Unstable	Unstable	Weak	Weak
LG	Weak	Weak	Weak	Weak	Strong	Weak	Weak	Weak	Weak
EQ	Weak			Weak	Weak			Weak	Weak
SMR									
State	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
DM									
MoE	Weak	Weak	Weak	Strong	Weak	Weak	Weak	Strong	Weak
UR	Strong	Unstable	Unstable	Strong	Weak	Unstable	Unstable	Weak	Weak
LG	Weak	Weak	Weak	Weak	Strong	Weak	Weak	Weak	Weak
EQ	Weak			Weak	Weak			Weak	Weak
SEQ									
State	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
DM									
MoE	Weak	Weak	Weak	Strong	Weak	Weak	Weak	Strong	Weak
UR	Strong	Unstable	Unstable	Strong	Weak	Unstable	Unstable	Weak	Weak
LG	Unstable	Weak	Weak	Weak	Strong	Weak	Weak	Weak	Weak
EQ				Weak	Weak			Weak	Weak

Table 4.5: Evolution of the Elmira Conflict.

	Status Quo	Transitional State	Transitional State	Cooperative Equilibrium
	Transition through state 6			
MoE: 1. Modify	N	N →	Y	Y
UR: 2. Delay	Y	Y	Y →	N
3. Accept	N	N	N →	Y
4. Abandon	N	N	N	N
LG: 5. Insist	N →	Y	Y	Y
State Number	**1**	**5**	**6**	**8**
	Transition through state 7			
	Status Quo	Transitional State	Transitional State	Cooperative Equilibrium
MoE: 1. Modify	N	N →	Y	Y
UR: 2. Delay	Y	Y	Y →	N
3. Accept	N	N	N →	Y
4. Abandon	N	N	N	N
LG: 5. Insist	N →	Y	Y	Y
State Number	**1**	**5**	**7**	**8**

In the foregoing situation, MoE have an incentive not to carry out its obligation and move from the intermediate state 7 to the final state 8 as state 7 is more preferred by both MoE and LG. On the other hand, UR can at any point in time move the entire conflict to an end point which is state 9. As state 9 is not favorable by both MoE and LG, they will try reaching a compromise which is state 8 in this case.

4.3 Summary

The foregoing analysis of the groundwater contamination conflict in Elmira demonstrates the effectiveness of the new fuzzy preference structure in dealing with a situation in which the DM's preferences are fuzzy or uncertain. The well-known solution concepts of Nash, general metarational, symmetric metarational, and sequential stability, were modified to accommodate preference fuzziness for multiple decision-maker conflicts. Despite the preferences' fuzziness, one needs to express his or her preferences over a set of states or course of actions hoping for a better or more rewarding strategic result. More specifically, it gives a more realistic way of analyzing the behavior of individuals involved in a conflict, especially when the preferences are vague or imprecise. Using this approach, whenever the preferences are not known for sure, they could vary within ranges rather than just being single crisp values. For the case where the cardinal preferences are not known, they can be approximated using linguistic variables. Hence, the methodology also allows a DM to consider both the ordinal and cardinal cases. In conclusion, using the new fuzzy approach along with the newly redefined solution concepts provides a more flexible framework for analyzing conflicts having imprecise preferences. In Chapter 5, a thorough review of trust models is carried out to reveal their capabilities and compare how and why they differ among different disciplines.

Chapter 5

Assessment of Trust Research

Trust decisions are risky due to the uncertainties and the loss of control. On the other hand, not trusting might mean giving up some potential benefits. The advances in electronic transactions, mutliagent systems, and decision support systems create a necessity to develop trust and reputation models. The development of such models will allow for trust reasoning and decisions to be made in situations with high risk and uncertainty. In recent years, several attempts have been made to model reputation and trust. However, perceiving trust differently and the lack of having a unified trust definition are among the main causes of the proliferation of many trust models across different disciplines. A thorough review of trust models is carried out in this chapter to reveal their key capabilities and compare how they differ among disciplines.

5.1 Overview

Despite its usefulness in both human and artificial societies, trust was and will still be a risky proposition.Trust has been always an integral component of human social life and actions. The advances in autonomous intelligent systems, communications, and electronic transactions motivated the execution of research regarding trust and reputation in order to span the spatial and temporal separation among the partners involved in a social interaction or an exchange of a commodity or goods.

Reputation is used as a means to build and update trust after a certain number of successful transactions (e-bay, 2002; Amazon, 2002; Dellarocas, 2003). Anonymity, uncertainty, risk, lack of control, and potential opportunism are key elements in most online transactions. Using trust evaluation and models to compensate for the lack of information and control in online environments will allow one to make decisions in regard to whom to trust and engage within a transaction or cooperative action. Some of the associated risks with online transactions include the exchange of some personal information, the absence of the physical goods, the non immediate exchange of goods and money, and how secure the transaction media is (Bhattacherjee, 2002; Einwiller et al., 2000; Einwiller and Will, 2001; Grabner-Krauter and Kaluscha, 2003).

Trust is closely related to and tied up with uncertainty. In general, uncertainty is the absence of credible knowledge about future events. Trust is supposed to assure an agent that the desirable course of events will be realized in the unknowable future as if being guaranteed from past knowledge. As Luhmann (1979, p. 32) wrote, “trust rests on illusion. In actuality, there is less information available than would be required to give assurance of success. The actor willingly surmounts this deficit of information”. A trustful person can comprehend new experiences and carry out actions that have been previously undesirable or unachievable. This is due to the fact that when trusting, in favour of an inner confidence, one simplifies the complexity of the outer world and removes any uncertainties.

When trusting, we allow ourselves to be vulnerable to others by depending on them to achieve or care for something we value. This interdependence relationship occurs when it is in the mutual benefit for both parties to fulfill their obligations towards the achievement of their common goal. In this case, no party has a dominant power over the other. While engaging in a dependence relationship, none of the parties is willing to exploit its situation. The realization of this dependency relationship by the trustee will put him or her in a relatively more powerful situation (Luhmann, 1979). If properly used, this kind of power will strengthen the trust relationship. Some trustees might refrain from using this power to avoid the negative consequences associated with exploiting the dependent trustor.

5.2 Trust Definitions from Different Disciplines

Different disciplines handle trust differently according to their own perceptions and what fits their specific goals. In order to consolidate sensible measures of trust, one needs to step back and analyze why different disciplines view trust differently. What follows is a thorough review of the existing trust definitions from different disciplines like, Psychology (Deutsch, 1973; Karlins and Abelson, 1970; Bromley, 1993; Rotter, 1967, 1971; Lewicki and Bunker, 1995), Sociology (Buskens, 1998; Luhmann, 1979; Lahno, 2001; Good, 2000; Sztompka, 1999; Gambetta, 2000), Philosophy (Plato, 1955; Hume, 1975; Hardin, 2002; Baier, 1986; Cvetkovivh and Lofstedt, 1999), Economics (Celenttani et al., 1966; Marimon et al., 2000; Granovetter, 1985; Zucker, 1986; Williamson, 1993), Finance (Guth, 2001; Ferrary, 2002), Marketing (Ganesan, 1994; Morgan and Hunt, 1994; Doney and Cannon, 1997; Geyskens et al., 1997; Swan et al., 1999), Management (Mayer et al., 1995; Gill and Butler, 1996; Inkpen and Currall, 1998; McKnight et al., 1998; Wicks et al., 1999; Luo, 2002), E-Commerce (Gefen, 2003, Gefen et al., 2003;

McKnight and Chervany, 2002; Jones, 2002), and Computer Science (Demolombe, 2001; Falcone and Castelfranci, 2001; Bhattacherjee, 2002; Shankar et al., 2002; Pavlou, 2003)

5.2.1 Trust in Psychology

In his 1973 book, M. Deutsch defines trust as confidence that one will find what is desired from another person rather than what is feared. Many researchers find this definition to be a specific characteristic of a relationship. Deutsch, however, presents many other aspects of trust in his 1973 work. He presents trust as being connected to despair, innocence, social conformity, virtue, gambling, risk-taking and faith, among others.

From a psychological perspective, risk in trust is approached as one of the characteristics of individuals. While some people are willing to take risks, there are others who are too cautious and distrustful to take any chances. Trusting behavior depends on how individuals perceive an ambiguous path or unclear situation. In such cases, the occurrence of a good or bad result is dependent on other's actions. Knowing that a negative result is more harmful than a good one, a trusting decision should be made.

The use of the word “perceive” in the previous paragraph is to emphasize the subjective nature of trust. If trust is based on individual perception, it is likely that the same situation will be seen differently by different individuals. Estimates of chances and expected gains or losses are subjective. Thus, some individuals might make unwise risks, thereby acting as if they are taking chances while, in fact, they are trusting unwisely.

5.2.2 Trust in Sociology

The sociology of trust has been investigated from different angles: rational choices, culture, functionality, symbolic interaction, and others. Trust is a social relationship subject to its own special system of rules (Luhmann, 1979). Trust occurs within interactions that are influenced by both personality and social systems (Lahno, 2001). Most sociologists agree with: "the clear and simple fact that, without trust, the everyday social life which we take for granted is simply not possible" (Good, 2000, p. 32). We always find ourselves in a condition of uncertainty about and uncontrollability of future actions. We have no way of knowing and controlling what others will do independently of our own actions and we are not even sure how they will react to ours. In general, uncertainty and risk are integral components of human interactions that can't be ignored or avoided.

In situations in which we have to act in spite of uncertainty and risk, the third factor that comes to the fore is that of trust (Sztompka, 1999). Trusting becomes a crucial strategy for dealing with an uncertain and uncontrollable future. Since there is no way of knowing what is in the minds of others, we need trust to deal with an unknown future and others' uncontrollable actions.

When participating in uncertain and uncontrollable conditions, we take risks, we gamble, and we make bets about the future and the actions of the others. A simple and general definition of trust is: "trust is a bet about the future contingent actions of others" (Sztompka, 1999, p. 25). In this sense, trust consists of two main components: beliefs and commitments. First, it involves specific expectations: "trust is based on an individual's theory as to how another person will perform on some future occasion" (Good, 2000, p.

33). When placing trust, we behave as if we know the future. Second, trust involves commitment through action or roughly speaking, placing a bet. Thus: “trust is the correct expectations about the actions of other people that have a bearing on one’s own choice of action when that action must be chosen before one can monitor the actions of those others” (Gambetta, 2000, p. 51). In order to have a better and deeper understanding of trust, we need to pay attention to the mental and subjective attitudes of the trusting person. It is important to focus on what happens in an individual’s mind when trusting someone else.

5.2.3 Trust in Philosophy

Trust and distrust are subjective attitudes that affect our thinking and feelings (Hardin, 2002). When trusting, we are more likely to let ourselves be vulnerable to others and allow ourselves to depend on others. Trust is a cooperative activity in which we engage so that we can assist one another in the care of goods (Baier, 1986). We trust others when we afford them the opportunity to care for something we value. We trust things as well as people. While trusting things is based on the properties of the things that we know in advance, trusting people is based on past experiences. When we trust, we hold expectations toward another person. To expect is to look forward to something without anticipating disappointment. When holding expectations of another, we project into the future, making an inference about the sort of person someone is going to be in the future. When trusting, the expectations alone are not enough but we must anticipate that the other has good intentions and the ability to carry out what is expected of him or her.

In order to trust someone, we need to have a sense of his or her values. A person who lacks commitment to any values or principles doesn’t give us the ability to predict either good or bad intentions or treatment. Knowing the other's values, commitments, and loyalty will help us to decide to what extent risk would be involved if we count on that person. We trust others more fully when we believe that they have positive feelings towards us personally and not just as members of some group. Trust is a risky business because people whom we trust can let us down and we are vulnerable to harm when they do so. It is important to accept the risks of trust and try to handle them rather than taking the simplistic view that trust is always good. Sometimes we trust too easily and risk a great deal in doing so (Cvetkovivh and Lofstedt, 1999; Hardin, 2002). Our trust is generally based on experiences with other people. On the basis of those experiences, we construct a characterization or picture of them but in reality they are free agents with different characterizations that go beyond our beliefs about them.

5.2.4 Trust in Economics

The study of trust and reputation in a free economy tries to address the relationship between trust and competition. By supplying quality goods at competitive prices, firms are building good reputations in order to secure their future market position and share. Firms will refrain from being concerned about the short-term profits when compared to building a good reputation and long-term profits thereafter.

In free-markets economy, consumers are faced with the dilemma of getting quality good for the least prices from profit-maximizing entities (firms). The trade-off between the price of goods and their quality is bridged by means of good reputation and trust between the consumer and the goods' providers. Some of the pioneers in this field are Shapiro (1983), Friedman (1960), Hayek (1974 and 1978), Rockoff (1975), Chang (1998), Chari and Kehoe (1990), Stokey (1991), Klein (1974), and Taub (1985, 1986).

5.2.5 Trust in Finance

The allocation of financial resources to certain activities includes buying assets, investments, and loans. These activities, and all financial activities, in general, are associated with some risk and uncertainty due to one of the involved parties not honoring his or her obligations. For example, borrowing money for a specific investment is highly related to the future ability of the borrower to pay back the loan. This highly depends on the trustworthiness and the associated evaluation of risks. Jensen and Meckling (1976) strongly believe that trust, reputation, and social bonds will always be present in such interactions. The formation of trust and what factors would affect it were a topic for research in finance. Hart (1987) studied trust within agency theory. Others, like Shapiro and Stiglitz (1984), investigated the ethical side of trust in terms of the reliability of one of the parties. This required importing some of the sociological concepts such as social capital and social networks (Granovetter, 1973, 1974, 1985, 1991). Guth and Kliemt (1994) analyzed the evolution of trust in a simple game of trust between a buyer and a seller.

5.2.6 Trust in Marketing

Studying trust relationship between a marketer and a customer is a key factor in the relationship between the two. Most of the research in this area focuses on the customer's trust (Morgan and Hunt, 1994). The research of trust in marketing dates back to the 1970s. Establishing a high level of trust in a marketing relationship allows the two parties to focus more on long lasting term benefits (Ganesan, 1994). Some of the developed marketing theories are based on trust (Morgan and Hunt, 1994). Trust could assume different phases like, the trust between the firm and its marketers, the marketers and the customers, and the customers and the firm. These three trust phases interact and affect each other one way or another (Morgan and Hunt, 1994). This explains why most marketing researchers have included trust in their relationships channel models where a vendor provides a service or a good to a distributor who resells it to the end user (Geyskens, 1997).

5.2.7 Trust in Management

Different parties within an organization need to work together to accomplish specific goals at both the personal and the organizational levels. This often requires some teamwork and dependence on others to execute certain assigned tasks. Risk will be always present in such relationships due to a lack of knowledge to do a specific task or the unwillingness to do it (Mayer, 1995). The presence of trust will reduce the risks associated with group interactions. However, some of the problems associated with trust in such environments are: lack of a specific definition of trust, difficulties of defining the boundaries of each task, lack of well defined regulations governing the interactions between the different inter-organization parties, and the unclear relationship between trust antecedents and consequences. There are some studies suggesting that trust is highly influenced by factors of which some are individual and others are organizational. In his 1998 work, Doney et al. suggested that social values and norms, besides behavioral attitudes, are key factors in trust. The length and the type of the relationship between the different parties within an organization and between the different organizations, the presence of previous interactions, and the interpersonal relations, if any, are other factors suggested by Inkpen and Currall (1997). Gill and Butler (1996) focused on the presence of some personal knowledge or quality for fulfilling some delegated tasks. Therefore, they define trust as an elaboration from current qualities as the most reliable for attaining a future goal. Some hidden factors or mental processes could be accounted for in explaining the high levels of trust for entities interacting for the first time (McKnight et al., 1998). Trust leads to some interdependencies which will eventually involve some sort of sharing of the control and management of things we care for (Inkpen and Currall, 1998). Nevertheless, trust has not been appreciated enough within the management field. This is in part because managers didn't devote sufficient time, energy, or resources to creating it within their organizations or because they look at it as a matter of strategic choice (Wicks et al., 1999).

5.2.8 Trust in E-Commerce

Trust in electronic transactions goes beyond risk and uncertainty to include other factors such as lack of information, lack of control, ease of use, privacy and security issues. On-line transactions and exchange relationships are not only characterized by uncertainty, but also by secrecy, lack of control and potential fraud, thereby making risk and trust crucial elements of electronic commerce.

The process of buying over the internet being perceived as risky, presents numerous risks for consumers during and after the transaction itself. Online firms may be located in different locations of the country or even in different countries. This requires a non-immediate exchange of information, goods, and money. As a result, some sensitive information is exchanged online like, personal and financial information. The limited history about the seller prior to the interaction adds to the risk and uncertainty involved in this transaction (Bhattacherjee, 2002; Gefen, 2003; Gefen et al., 2003).

Some of the system-dependent uncertainties go beyond the control of the parties involved in the transaction. These are environment related uncertainties which could be characterized as exogenous. Generally speaking, the concept of exogenous uncertainties refers to the uncertainties of the world (Hirshleifer and Riley, 1979). The environment dynamics and system complexity are two main factors when considering exogenous uncertainties (Brielmaier and Diller, 1995). In the context of electronic commerce, exogenous uncertainty relates to the potential technological errors or security gaps that can’t be avoided. The utilization of encrypted transactions, firewalls, authentication mechanisms and privacy seals are means of reducing the effects of system uncertainties (Pavlou, 2003). Transaction-specific uncertainties are caused by decisions of parties exchanging information over the transaction media (Weiber and Adler, 1995b). The consumer may interpret the uncertainties as seller’s potential behavior in the transaction process. In computer mediated transactions, element of personal interaction like body language, gestures, and facial expressions are eliminated (Winand and Pohl, 2000).

In general, the more trust present in a given situation, the less additional information is needed to make a certain decision. On the other hand, if there is little or no trust, there will be a need for complete information in order to reduce system-dependent and transaction-specific uncertainties. Uncertainties are perceived differently and, hence, the level of the perceived uncertainties influences the needed balance between trust and information (Tomkins, 2001). Trust and additional information could be seen as means to reduce uncertainties (Luhmann, 1989; Wicks et al., 1999).

5.2.9 Trust in Computer Science and Information Systems

Computer scientists tried to formalize the measures of knowledge derived from sociology and psychology into agents' architectures. One can understand trust as an attitude of an agent who believes that another agent has a given property. Therefore, one can analyze the meaning of trust as a function of the attributed properties. For instance, the property may be that the agent one trusts fulfills his obligations, like the case of a buying agent. Properties one considers are the ability of the agent to do the job, to make decisions, or just to deliver information (Demolombe, 2001; Gefen, 2003; Gefen et al., 2003).

With the emergence of electronic commerce, trust issues became important for many people. Generally speaking, it is agreed that in order for electronic commerce to become successful, most people have to trust it. The person's trust in a transaction is determined by the trust in the counter party and the trust in the transaction media based on the assumption that party and media trust supplement each other. If there is not sufficient party trust, then the media trust and its control protocols should be brought in to supplement the party trust. Trust in the counter party can be defined as "The subjective probability by which an individual A expects that another individual B performs a given action on which its welfare depends" (Falcone and Castelfranci, 2001, p. 56). According to this definition, it could be argued that trust has both objective and subjective attributes. The first depends on the media structure, such as the functionality of the control mechanisms in place. The second depends on personal experiences in dealing with a specific party, or with specific procedures and control protocols.

5.2.10 Trust as a Global View

Gambetta (2000) attempted to gather different thoughts regarding trust from many areas. The most important aspect of his work is the use of values. On the other hand, using explicit values for trust can be problematic due to the subjectivity of trust in which the same value could be seen differently by different agents. Yet the use of values for measuring trust allows one to talk more precisely about certain circumstances or behaviors concerning trust. Also, it permits a straightforward implementation of the formulation.

In his research, Gambetta (2000, p. 217) defines trust as "a particular level of subjective probability with which an agent assesses that another agent or group of agents will perform a particular action both before he can monitor such action or independently of his capacity ever to be able to monitor it and in a context in which it affects his own action". This definition excludes certain aspects which are important to trust like referring only to the trust relationship between the agents themselves and not, for example, the agents and the environment. It also excludes those agents whose actions have no effect on the decision of the truster, despite the fact that trust is present. An interesting point in Gambetta's work is the concern regarding competition and cooperation. In some cases, cooperation is not good, such as the cooperation among thieves or drug dealers, while it is very desirable among policemen. Then, it is beneficial to find "the optimal mixture of cooperation and competition rather than deciding at which extreme to converge" (Gambetta 2000, p. 215). In competitive situations, cooperation is of great importance since "even to compete in a mutually non-destructive way one needs at some level to trust one's competitors to comply with certain rules" (Gambetta 2000, p.

215). Despite the importance of using values for trust, Gambetta didn't develop the idea in any concrete fashion (Marsh 1994).

5.3 Approaches to Modeling Trust

Different approaches have been used in an attempt to model trust, of which some have commercial applications and others are only meant for academic purposes. Some of these modeling attempts are only informative while others are conceptual. In the following sections, different approaches for modeling trust are classified based on their underlying methodologies.

5.3.1 Simple Scoring

Considered a relatively simple approach, some basic mathematical operators like, multiplication and addition, are used to compute trust values. The average and the weighted average are the two most common methods in this category. Getting direct ranking or feedback from the users and then averaging all the responses is a simple and intuitive way of the many techniques used in e-commerce (Amazon, 2002). A slightly modified version of this technique is being used in e-Bay (e-Bay, 2002). Both positive and negative scores are summed separately and then subtracting the total negatives from the total positives to get the overall score. The values often used are 1, 0, and -1 for the positive, neutral, and negative ratings, respectively. In some cases, the weighted average is being implemented to put more emphases on the most recent transactions or to highlight some factors more than others.

5.3.2 Statistical

When using this technique, a history of all previous interactions is maintained. This history is combined with the new interactions to compute the overall trust value using statistical approaches. The most common approach is Bayesian. The Bayesian system takes a binary input and utilizes the beta-Probability Density Function (PDF) to compute the updating. Within the PDF distribution, the two parameters (α, β) refer to the positive and negative ratings, respectively.

The Bayesian system starts with 1 assigned to both parameters and keeps updating after each interaction. While this provides a sound theoretical basis for computing a trust value, it might not be easily understood by average users (Josang, 1999; Mui and Mohtashemi, 2001; Josang and Ismail, 2002; Mui et al., 2002).

5.3.3 Linguistic

Sometimes, it is easier describing the level of trust using some linguistic terms rather than numerical values. Using fuzzy or probabilistic approaches, those linguistic terms could be matched with appropriate or approximated numerical values that are easy to calculate and program. Al-Mutairi et al. (2005b) used the linguistic terms absolutely low, very low, low, fairly low, medium, fairly high, high, very high, and absolutely high to describe the trust level. This enables the agent to calculate the trustworthiness of another agent before engaging with it in an interaction. Fuzzy logic is used to match those linguistic terms with approximated values to carry on some computations and obtain the overall expected trust value. This also depends on some other factors like the importance of the interaction for a specific agent, the expected value, the availability of other alternatives, and the risk attitudes of the agent.

5.3.4 Cognitive

This technique tries to mimic the human way of thinking and reasoning about trust. It attempts to go beyond sensible things and explore what transpires in the mind of one when trusting. This is highly linked to one's belief and social community. For an inner feeling or confidence one may or may not trust another person. The thresholds of what is trustworthy or not will be different for different agents. Some authors (Josang, 2001) use the belief theory to predict a trust value. Belief theory is a framework based on probability theory where the total of the probabilities doesn't necessarily add up to 1. This is in part due to the presence of some uncertainties. It is important to mention that transitivity is an underlying assumption in most of the models in this category where an agent is considered as trustworthy if referred to as trustworthy by other agent or agents.

5.3.5 Fuzzy

When using fuzzy logic to evaluate trust, it is possible to refer to trust using a linguistic label that describes a specific fuzzy function rather than using numerical values. The trust level can have different memberships to different fuzzy sets like belonging to trustworthy and very trustworthy with memberships of 0.4 and 0.6, respectively. The models proposed by Al-Mutairi et al. (2005b), Manchala (1998), and Sabater and Sierra (2001, 2002) are good examples of this type of modeling.

5.3.6 Flow Chains

The main assumption underlying this category of models is transitivity. By that, one means that if agent a trusts agent b, agent b trusts agent c, then agent a must trust agent c. It could be as simple as an interaction between three agents or through long chains and loops of iterative deals. However, it could be the case that trust values from different agents are assigned different weights depending on the previous history of that particular agent. More interaction chains through a particular agent means higher trust value and vice versa. In web semantics, the more hyperlinks to a site the higher its rank and more hyperlinks out of that page the less its rank (Page et al., 1998).

5.4 Discussion

Table 5.1 shows a chronological summary of some of the existing work in trust modeling. From this extensive review, one can highlight the following issues for further investigation when modeling trust.

5.4.1 Unrated Transactions

Though sighted as one of the most common ways of evaluating the rules of trust, feedback is not always given for all transactions (Resnick and Zeckhauser, 2002). This is in part due to the following:

- Lack of incentive (no direct benefits of providing feedback).
- Retaliation from the seller or service provider in response to negative feedback.
- Competition for a limited service or commodity.
- Feedback mechanism is lengthy or not easy to use.
- Ignorance.

Thinking of feedback as only being important in case it is negative (a way of warning others) while neglecting the positive ones could give a misleading trust value. For example, on e-Bay, assume that only 50 out 1000 transactions are assigned a negative feedback (the remaining 950 are positive). The score will be 950 when all transactions are given a feedback while the score for the same seller will be 750 if only 800 out of the 950 positive transactions are reported. This will cause the positive feedback ranking for this seller to drop from 0.95% in the first case to only 0.75% in the second case.

Table 5.1: Chronological summary of some of the existing trust models.

Year	Author(s)	Domain	Methodology	Remarks
1994	S. Marsh	Computer Science	Simple Scoring	Simple mathematical formulation for multiagent systems.
1998	C. Castelfranchi and R. Falcone	Multiagent Systems	Cognitive	Based on goals, mental states, and beliefs.
1998	D. Manchala	Electronic Commerce	Fuzzy	Focuses on the relationship between trust and risk
1998	L. Page et al.	Electronic Commerce	Flow Chains	Transitivity through loops or long chains of interactions
1999	G. Zacharia	Electronic Commerce	Simple Scoring	Sporas and Histos are two modified models for online reputation systems with a focus on recent ratings.
1999	A. Josang	Electronic Commerce	Statistical	Based on statistically updating Beta probability density functions
2000	M. Schillo et al.	Multiagent Systems	Statistical	Boolean logic where it is either strictly good or bad.
2000	A. Abdul-Rahman and S. Hailes	Multiagent Systems	Linguistic	Based on witness information with some adjustments
2000	J. Schneider et al.	Electronic Commerce	Simple Scoring	Averaging all ratings (both positive and negative)
2001	B. Esfandiary and S. Chandrasekharan	Multiagent Systems	Statistical	Trust acquisition using Bayesian learning
2001	L. Mui and M. Mohtashemi	Electronic Commerce	Statistical	Based on statistically updating Beta probability density functions
2001	A. Josang	Electronic Commerce	Cognitive	Based on belief theory where the probabilities don’t necessarily add up to 1.
2001,2002	B. Yu and M. Singh	Multiagent Systems	Statistical	Only most recent information is considered for calculation
2001,2002	J. Sabater and C. Seirra	Electronic Commerce	Fuzzy	Aggregated information from direct and indirect interactions.

Table 5.1: (*continued*)

2002	S. Sen and N. Sajja	Multiagent Systems	Statistical	Both direct interactions and observed ones are considered.
2002	eBay, Amazon, OnSale	Electronic Commerce	Simple Scoring	Online reputation models through direct feedback.
2002	J. Carbo et al.	Multiagent Systems	Fuzzy	Uses weighted aggregation to combine old and new reputation values.
2002	J. Carter et al.	Multiagent Systems	Cognitive	Uses weighted aggregation but the values used are different for different societies.
2002	A. Josang and R. Ismail	Electronic Commerce	Statistical	Based on statistically updating Beta probability density functions
2002	L. Mui et al.	Electronic Commerce	Statistical	Based on statistically updating Beta probability density functions
2003	V. Cahill et al.	Electronic Commerce	Linguistic	Some heuristics are needed to associate linguistic labels to values
2003	M. Carbone et al.	Electronic Commerce	Linguistic	Some heuristics are needed to associate linguistic labels to values
2003	S. Kamvar	Electronic Commerce	Flow Chains	Transitivity through loops or long chains of interactions
2004	A. Withby	Electronic Commerce	Statistical	Based on statistically updating Beta probability density functions
2004	C. Zeigler	Electronic Commerce	Flow Chains	Transitivity through loops or long chains of interactions
2004	R. Levien	Electronic Commerce	Flow Chains	Transitivity through loops or long chains of interactions
2005	E. Maximilien and M. Singh	Multiagent Systems	Statistical	Aggregate different scores for multiple attributes and choose the agent with the highest score
2006	N. Griffiths et al.	Peer-to-Peer Systems	Fuzzy	Combine a set of rules to represent and reason about others' trustworthiness

5.4.2 Misleading Feedback

Feedback could be misleading when, for some reason, it is unfair or not justified whether they are positive or negative. Some of the sited reasons for having a false positive feedback are:

- Reciprocation: a positive feedback for a positive feedback in return.
- Being rewarded with a discounted price.
- Building a good reputation through prearranged fake interactions.

In contrast, false negative feedback could be due to:

- Based on a specific identity of a specific agent whether it is because of a previous interaction history or personal reasons.
- Blaming the seller or the service provider for a shortcoming on behalf of the buyer or the service recipient.
- Reasons that are beyond the control of the seller which could be related to the transaction media or the delivery system.

The process of providing feedback is a very subjective issue that is hard to monitor and control (Miller et al., 2002).

5.4.3 Identity Verification

One of the risks associated with electronic environments is verifying that an agent is what he or she is claiming to be (Zacharia et al., 1999). Some of the identity associated risks are:

- Stolen verification information (username and passwords).
- Identity change to escape from a past transaction history.
- Validating the information supplied during the registration process.

Based on the assumption that trust is the result of acquired cumulative reputation over a period of time through a number of interactions, not being able to verify the agents' identities will give a misleading trust index (Pavlou, 2003).

5.4.4 Behavioral Changes

When showing good intentions, regardless of the current low trust index, agents need to be given a chance to recover and start a corrective process (Jøsang and Ismail, 2002). An agent might start with a low trust index for one or more of the following reasons:

- Focusing on the short term benefits and not worrying about a long lasting one.
- Lack of knowledge about the importance of building a good reputation on the virtual environments.
- Reasons that are related to the transaction media which is beyond the control of the agent.
- Change of the service provider' management in order to recover from the current situation.
- Change of the service type or product.
- Behavioral changes over time.

Giving more weight on the most recent transactions, like for the past six months or last year without entirely neglecting the past interactions (Buchegger et al., 2003a,b), will give a more reflective index of the agent's current situation. This will also allow one to analyze any behavioral trends over a period of time.

5.5 Summary

From this extensive review, one can appreciate the importance of trust across many disciplines. However, trust research is still in its early stages and varies greatly depending on the trust context and use. Most of the models are based on feedback through direct interactions or conveyed through a third party. Though agreeing on the importance of direct experiences, there are more factors that contribute to trust that should be taken into consideration. By its nature, trust is complex, multidimensional, and subjective. It might be time to merge traditional game theoretic approaches with cognitive, sociologists, and psychological ones in order to better understand and model trust. Due to the variation in defining and using trust, as of now, there is no single set of unified trust data that could be tested and compared among the different trust models. Testing and comparing trust models are still an arbitrary issue. Developing test data sets and general test frameworks will enable fine tuning and improving some of the proposed models. It will also enable researchers to examine which model works better for which uses. Chapter 6 presents the newly developed fuzzy trust modeling.

Chapter 6

Modeling Trust using a Fuzzy Logic Approach

In open and unpredictable environments, one might have incomplete, misleading, vague, imprecise, or ambiguous information. Since fuzzy logic has been demonstrated to work well in such environments, fuzzy logic will used to model trust. This will be of great importance in the areas of decision making, e-commerce, and multiagent systems in order to capture some of the complexity and dynamics of trust. The proposed model doesn't question the validity of the other existing models but rather tries to bring them all together in one robust model that capture all the important elements of trust.

6.1 Overview

In order to build a long term relationship to achieve some goals and benefits that are not achievable on individual basis, one needs to trust others and cooperate with them. This might require scarifying some short term benefits. To minimize the chances of loss and betrayal, one needs to make a trusting decision. Currently, trust is being modelled as a variable with a threshold for action. When the value of the trust variable exceeds the specified threshold, a binary decision is taken. This action in most of the cases is "cooperate" or "don't cooperate" and "trust" or "don't trust". Trust is only required in situations of risk and uncertainty. The trustor needs to decide between two alternatives; to "trust" or "not to trust". In a situation that requires trust, the potential loss is higher than the potential gain; otherwise the decision would not be one between trusting and not trusting but based on a rational "loss-benefit" analysis. The general framework for the proposed trust model is displayed in Figure 6.1. It consists of three main modules: the decision maker, the trustworthy assessment, and the fuzzy evaluation.

6.2 Decision Maker

Besides having positive feelings, to be able to make a trusting decision, we need to consider some objective factors from within the surrounding environment. A decision maker tries to collect as much information as possible about the person or agent of concern. This could come from a third party, past history (reputation), or by observation. A set of required or predefined factors to be known prior to making a trust decision include (but not limited to) information, actions and statements, exception clauses, circumstances, and judgment.

6.2.1 Information

A key factor in the trusting decision is how much knowledge one has about the trustee. The more information one has the better one understands the situation and estimates the possible gains and losses. Information could be acquired through:

1. Direct or personal knowledge of the potential trustee.
2. Indirect knowledge (through a third party).
3. Past experiences (reputation) whether it is direct or indirect.

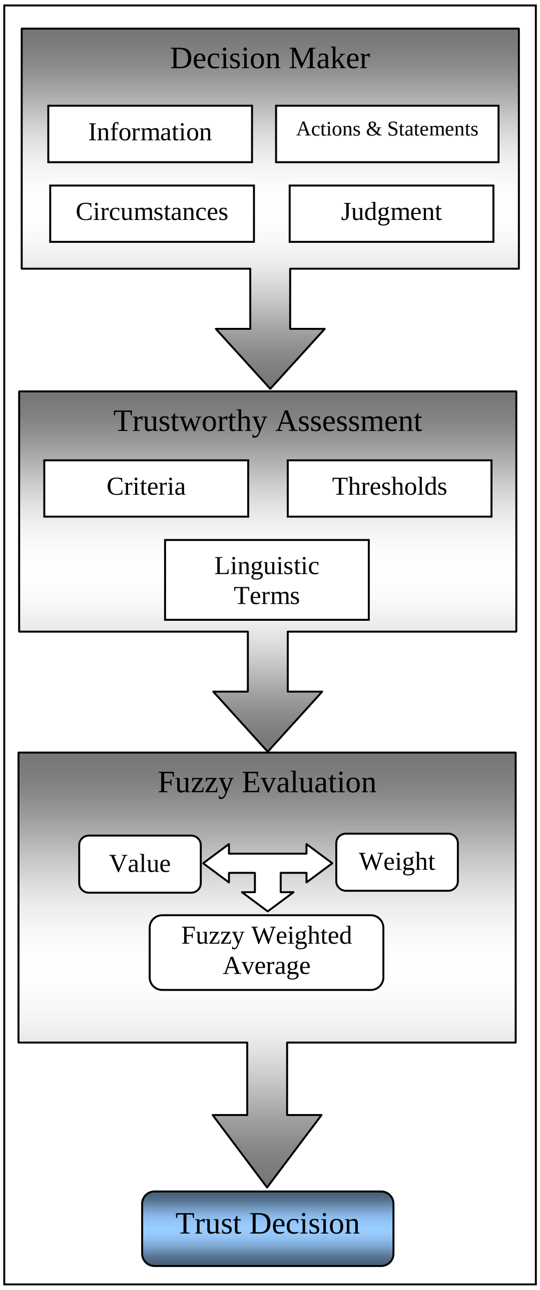

Figure 6.1: Fuzzy trust model framework.

4. Social Role (the position or the job of someone might give an indication of how much trust one can put on him or her).

Information from the above sources can be collected, interpreted, and evaluated to construct a characterization or picture of the persons or agents of concern but in reality they are free agents with different characterizations that go beyond one's beliefs about them.

6.2.2 Risk

When introducing the concept of trust one is already referring to risk whether it is in a direct or indirect way. Risk is activated by our actions, the choices we make, and the decisions we take (see Figure 6.2). When trusting, one is actually taking risks (accepting

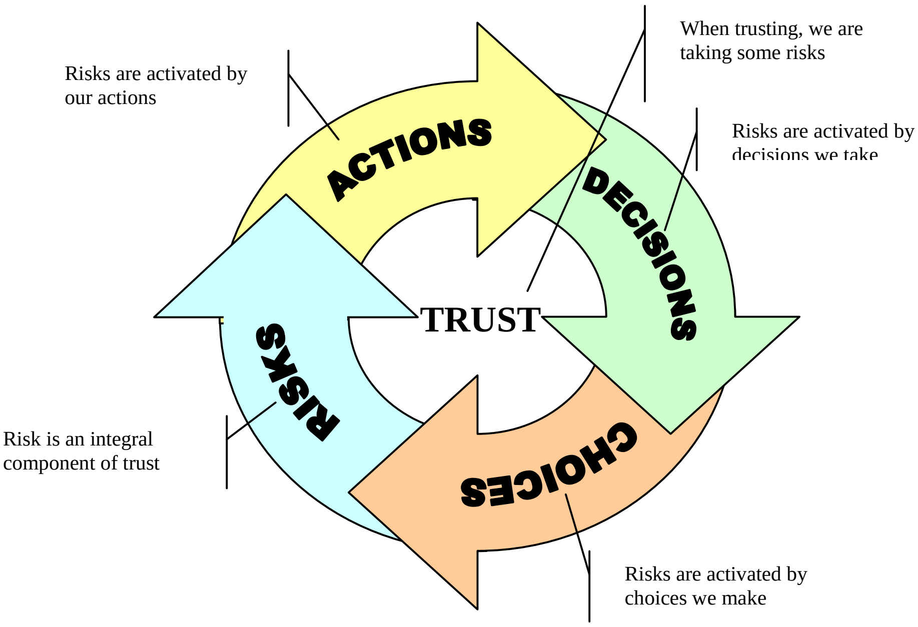

Figure 6.2: Risk and trust relationship.

the risks and trying to handle them). In other words, one tries to know what is going to happen in the future and ignoring the fact that risk is still there. To keep relationships going, it is necessary to take some risks. Depending on how one handles them, they might strengthen the existing trust or reduce it dramatically. Even if the trust decision is made based on evidence, there is always a chance of risk (Cvetkovivh and Lofstedt, 1999). Despite its usefulness in both human and artificial societies, trust was and will still be a risky proposition. Since the early years of studies about trust, risk has been considered one of the closest elements to the trust concept and one of the most important factors that affect trust decisions. It becomes even more typical and understandable that decisions cannot avoid risk (Zeckauser and Viscusi, 1990).

Based on the fact that risks have been always an integral component of decisions and actions, it could be claimed that they do not exist by themselves. If you refrain from action, you will not have any risk. Trust is based on a relationship between risk and action. Action defines itself in relation to a particular risk as a future possibility. Whether or not one places trust in future events, the perception and evaluation of risk is a highly subjective matter. The same situation will be perceived differently by different people depending on their personal attitudes and whether they are risk-seekers or risk-averse.

It could be argued that the decision whether to trust or not is a function of the expected gain and loss involved. The decision to place trust is similar to the decision to place a bet. A rational decision maker will place trust if the chance of winning, relative to the chance of losing, is greater than the amount that would be lost relative to the amount that would be won (see Figure 6.3). The sensible thing to do for the potential trustor is to collect as much additional information as possible on the potential gain and loss involved and the trustworthiness of the trustee. Information will have the effect of changing one's estimate of the probability of gain and the trust decision when compared to the predefined thresholds for making the decision.

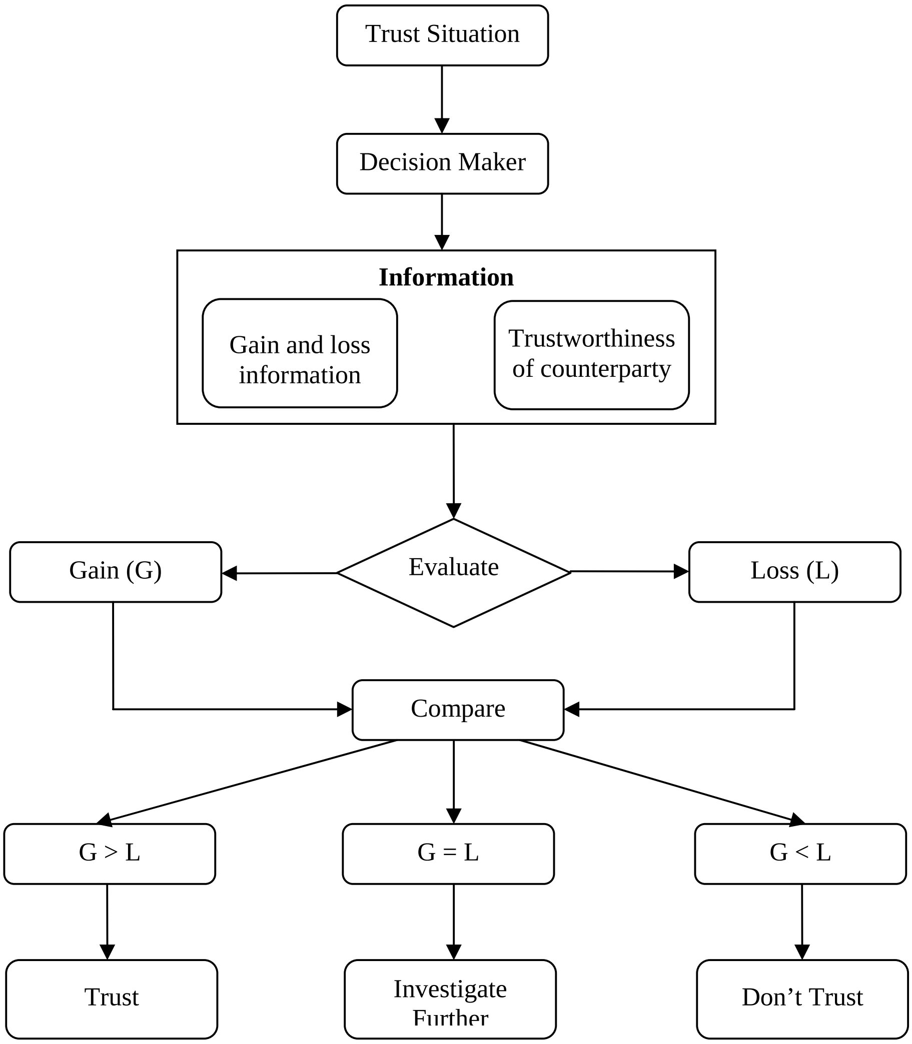

Figure 6.3: Risk evaluation process.

6.2.3 Uncertainty

Trust is closely related and tied up with uncertainty. In general, uncertainty is the absence of credible knowledge about future events. Trust is supposed to assure an agent that the desirable course of events will be realized in the unknowable future as if being guaranteed from the past knowledge. As Luhmann (1979, p. 32) wrote, “trust rests on illusion. In actuality, there is less information available than would be required to give assurance of success. The actor willingly surmounts this deficit of information”.

In favor of an inner confidence, a person in this illusory state of trust simplifies the complexity of the outer world and removes any external uncertainty. As a result of lack of information and a deceptive sense, the trustful individual finds a possibility of comprehending new experiences and carrying out actions that have been previously undesirable or unachievable.

The existing objective events and states of the outer world help to form an agent’s subjective assumptions, even if it is partial or unreliable, about whether or not to trust. Yet, the objective available evidence under consideration serves only as hints at the possibility of trust rather than ensuring its certainty. It is true that the state of partial knowledge (incomplete or unreliable) is the only feasible environment for emerging trust. In the extreme case of total knowledge or total ignorance, trust becomes an empty concept. With complete certainty, there will be no need or even possibility for trust to develop. On the other hand, in the case of absolute ignorance, there can be no reason to trust. When faced by the totally unknown, we can gamble but we cannot trust.

6.2.4 Dependence

When trusting, one allows him or herself to be vulnerable to others by depending on them to achieve or care for something one values. The interdependence relationship occurs when it is in the mutual benefit for both parties to fulfill their obligations towards the achievement of their common goal. In this case, no party is having a dominant power over the other. While engaging in a dependence relationship, none of the parties is willing to exploit its situation. The realization of this dependency relation by the trustee will put him or her in a relatively more powerful situation (Luhmann, 1979). If properly used, this kind of power will strengthen the trust relationship. Some trustee might refrain from using this power to avoid the negative consequences associated with exploiting the dependent trustor.

6.2.5 Reputation

Reputation is looking back at the history of the person and his or her past experiences and trying to predict how he or she would act in the future (Sztompka, 1999). It is more of a system with some feedback since after each interaction, it would improve (positive) or decrease his or her reputation (negative) based on his or her behavior in that specific interaction (see Figure 6.4).

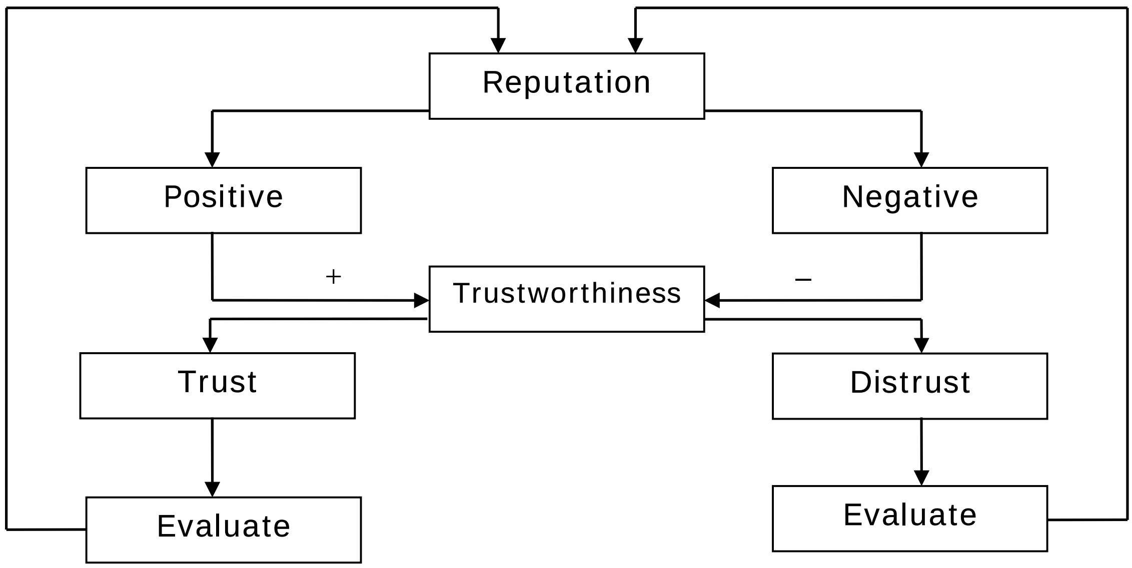

Figure 6.4: Trust and reputation.

Decisions about whether or not one should trust another person depend on that person's reputation, and a favorable reputation is something which everyone would like to establish (Zacharia and Maes, 2000). Therefore, the ways in which a good reputation is established or destroyed are important. Not only will the perceivers of a reputation usually have access to information which the reputation holder does not control, but also the manner in which the information are interpreted is not straightforward. Such information will often be ambiguous and open to many readings, and, even when it is not, it is not necessarily the case that all people will assign any particular item the same significance. The interpretation of a reputation will be a function of the interpreter and, therefore, it is necessary to consider how new information is handled and is related to existing views (Zacharia and Maes, 2000).

One usually wishes to know the sort of person one is dealing with before one deals with him or her. But one will know it only imperfectly. One forms an opinion on the basis of his or her background, the opportunities he or she has faced, the courses of action he or she has taken, and so forth (Sztompka, 1999). Our opinion is thus based partly on the theory one holds of the effect of culture, social class membership, and on a person's motivation and hence his or her behavior. The opinion which is publicly formed and held is this person's reputation. The problem, in essence, is to infer the person's qualities from such data.

For trust to be developed between individuals they must have repeated encounters, and they must have some memory of previous encounters. And finally, trust is linked with reputation, and reputation has to be acquired. Reputation is like a capital asset. One can build it up by pursuing certain courses of action, or destroy it by pursuing certain others. A reputation for honesty, or trustworthiness, is usually acquired gradually. Although a reputation for honesty may be acquired slowly, it can generally be destroyed very quickly (Hardin, 2002).

6.2.6 Actions and Statements

In the absence of credible information, one tries to construct a characterization or picture of someone based on his or her actions or statements. While truth telling, sincerity, promise keeping, keeping confidences, reliability, keeping appointments, and concern for others are signs for positive moral values and attitudes, lying, dishonesty, unreliability, and not keeping promises, on the other hand, are signs that we shouldn't invest much in this person. An overall sense of someone is derived from specific actions and statements that are projected to circumstances in future of some interest to the trustor. One should always keep in mind that these statements and actions might be specific to a certain person in a certain situation in a certain time.

6.2.7 Exception Clauses

Realising the fact that things might mean something different from what it appears to, it is wise not to generalize judgments without knowing the motives and circumstances. The use of exception clause is very important when interpreting someone else's actions. Someone's behavior might not reflect really what it appears to mean. For example, when I lock myself out and try to get in, for someone just passing by, I am just a thief but in reality I am just trying to get in my house. My action in this case seems untrustworthy but the reality is not what it looks to be.

6.2.8 Circumstances

Different circumstances require different degrees of trust or distrust. For example, to trust someone to help us carry some packages across the street is not a big issue assuming that the packages don't contain valuable items because the worst thing that could happen is that the person would run away with the packages. On the other hand, we should think twice before accepting assistance from someone we don't know well before accepting his or her help to transfer some valuable items. Finally, deciding to trust or distrust someone is like belief. One can't believe and disbelieve in something at the same time. To accept something and belief in it, one needs some kind of evidence and reasons for that belief.

6.2.9 Judgment

Based on how much information one knows about another one, how reliable the information, how to interpret his or her statements and actions, is there a chance that one misunderstood or misinterpreted his or her statements and actions, and what are the circumstances that surrounded those actions, one could make an overall judgment whether that person deserves to be trusted or not. Of course, when deciding to go for the extremes (totally trusting or totally distrusting) someone, one needs to look for solid evidence, whether they are positive or negative. There is little or no room for an exception or unless clauses in this case. In cases of little or limited trust or distrust, one may consider the exception clauses and consider some questionable information.

6.3 Trustworthy Assessment

A logical and distinct step separating trusting from gambling is the assessment of the trustworthiness of the agent or person of concern. The decision maker collects information, evaluates, and interprets them to make an estimate of the qualities of the agent of concern. By comparing those qualities with a predefined scale for the minimum acceptable levels, the trusting decision is to be made.

6.3.1 Trust Criteria

Based on a comprehensive and thorough study of trust among different disciplines as described in chapter 5, the three main criteria for trust are the available information about the agent of concern, the circumstances for that issue in regard to that specific agent, and the judgment of the agent in charge of making the trust decision. Each of those criteria are classified further and split into sub-criteria as shown in Figure 6.5.

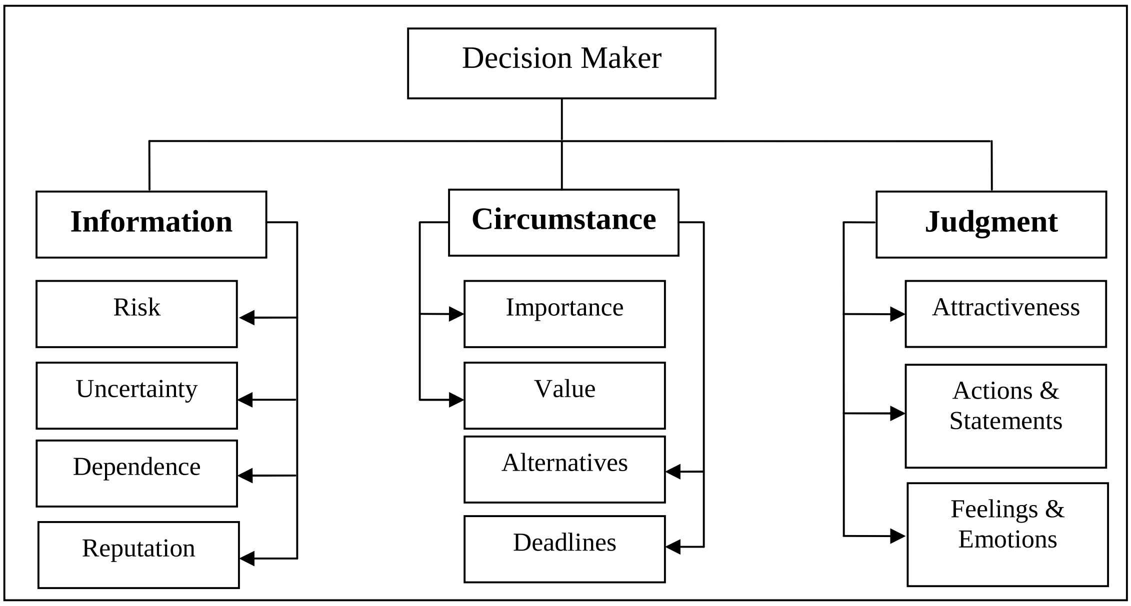

Figure 6.5: Trustworthy criteria and sub-criteria.

6.3.2 Fuzzy Values and Weights

The issue of trust is very subjective. Even trying to express it in natural language is not an easy task. The same value might be interpreted differently by different agents. Compared to crisp values, the use of natural linguistic terms is easier to understand and interpret and at the same time can handle some uncertainties. Fuzzy logic approach tries to mimic how humans would make their decisions. It starts with natural language concepts like, this person is trustful or I am willing to take some risk in my decision. It represents the values over an interval rather than a crisp value. It is good for modeling subjective issues like trust that is difficult to estimate experimentally. Another advantage is that it works well with vague, ambiguous, imprecise, noisy, or missing input information.

Due to the subjective nature of the evaluation criteria as well as the vague and imprecise nature of the available information, it is easier to express the values and the weights in natural linguistic terms rather than specifying crisp values. These linguistic terms could be assessed through the use of fuzzy logic. A six and five level scale weights and values and their corresponding fuzzy representations given in Tables 6.1 and 6.2, respectively.

Table 6.1: Weights and their fuzzy representation.

Linguistic Terms	Generalized Fuzzy Numbers
Totally Unimportant	(0, 0, 0.1, 0.1; 1.0)
Unimportant	(0.1, 0.2, 0.3, 0.4; 1.0)
Fairly Unimportant	(0.3, 0.4, 0.5, 0.6; 1.0)
Fairly Important	(0.5, 0.6, 0.7, 0.8; 1.0)
Important	(0.7, 0.8, 0.9, 1.0; 1.0)
Very Important	(0.9, 1.0, 1.0, 1.0; 1.0)

Table 6.2: Values and their fuzzy representation.

Linguistic Terms	Generalized Fuzzy Numbers
Low	(0, 0.1, 0.2, 0.3; 1.0)
Fairly low	(0.2, 0.3, 0.4, 0.5; 1.0)
Medium	(0.4, 0.5, 0.6, 0.7; 1.0)
Fairly high	(0.6, 0.7, 0.8, 0.9; 1.0)
High	(0.8, 0.9, 1.0, 1.0; 1.0)

6.3.3 Thresholds

As mentioned earlier, trust is subjective and multidimensional to the extent that it is difficult to design global thresholds that fit every situation. Such issue is relative to the agent in regard to a specific situation in specific time with specific attributes.

6.3.4 Linguistic Scaling

All criteria and thresholds are to be expressed using natural linguistic terms due to subjectivity and difficulty estimating them experimentally or in real life situations. It is up to the decision maker to decide on the importance and value of each factor. Converting these linguistic terms into numerical values is better accomplished through the use of fuzzy logic. Many popular linguistic terms and fuzzy values conversion scales have been proposed (Chen and Hwang, 1992; Karwowski and Mital, 1986; Miller, 1965). To get better and more accurate descriptive values and at the same reduce the unnecessary overlapping, it is recommended that the scales be between three and nine (Lin and Chen, 2004).

6.4 Trust Computation

After gathering all the necessary information and deciding on the trust criteria thresholds, a fuzzy computation is to be carried out to get the Fuzzy Weighted Average (FWA). Prior to that, a quick review of some of the fuzzy arithmetic will be covered in the following section.

6.4.1 Fuzzy Numbers and Their Arithmetic Operations

As stated by Chen and Chen (2003), a generalized trapezoidal fuzzy number is represented as $A = (a,b,c,d;w)$, where $0 \le w \le 1$, and a,b,c, and d are real numbers. If $w = 1$, then the generalized fuzzy number A is called a normal trapezoidal fuzzy number denoted $A = (a,b,c,d)$. If $a = b$ and $c = d$, then A is called a crisp interval. If $b = c$, then A is called a generalized triangular fuzzy number. If $a = b = c = d$ and $w = 1$, then A is called a real number.

Figure 6.6 shows a representation for a generalized trapezoidal fuzzy set $A = (a,b,c,d;w_1)$ which denote a fuzzy set describing a certain situation. The value of w_1 represents the degree of confidence of the situation of interest belonging to that group.

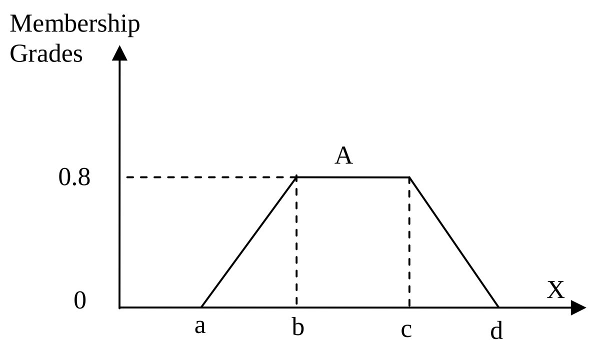

Figure 6.6: A generalized trapezoidal fuzzy number.

The arithmetic operations between the generalized trapezoidal fuzzy numbers A_1 and A_2 as follows:

1. Fuzzy numbers addition:

$$A_1 \oplus A_2 = (a_1,b_1,c_1,d_1;w_1) \oplus (a_2,b_2,c_2,d_2;w_2)$$

$= (a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2; \min(w_1, w_2))$ 6.1

Where $a_1, b_1, c_1, d_1, a_2, b_2, c_2,$ and d_2 are any real numbers.

2. Fuzzy numbers subtraction:

$A_1 \ominus A_2 = (a_1, b_1, c_1, d_1; w_1) \ominus (a_2, b_2, c_2, d_2; w_2)$

$= (a_1 - d_2, b_1 - c_2, c_1 - b_2, d_1 - a_2; \min(w_1, w_2))$ 6.2

Where $a_1, b_1, c_1, d_1, a_2, b_2, c_2,$ and d_2 are any real numbers.

3. Fuzzy numbers multiplication:

$A_1 \otimes A_2 = (a, b, c, d; \min(w_1, w_2))$ 6.3

Where $a = \min(a_1 \times a_2, a_1 \times d_2, d_1 \times a_2, d_1 \times d_2)$,

$b = \min(b_1 \times b_2, b_1 \times c_2, c_1 \times b_2, c_1 \times c_2)$,

$c = \max(b_1 \times b_2, b_1 \times c_2, c_1 \times b_2, c_1 \times c_2)$,

and $d = \max(a_1 \times a_2, a_1 \times d_2, d_1 \times a_2, d_1 \times d_2)$.

it is obvious that if $a_1, b_1, c_1, d_1, a_2, b_2, c_2,$ and d_2 are all positive real numbers, then

$A_1 \otimes A_2 = (a_1 \times a_2, b_1 \times b_2, c_1 \times c_2, d_1 \times d_2; \min(w_1, w_2))$.

4. Fuzzy numbers division:

The inverse of the fuzzy number A_1 is $\frac{1}{A_1} = (\frac{1}{d_1}, \frac{1}{c_1}, \frac{1}{b_1}, \frac{1}{a_1}; w_1)$ where $a_1, b_1, c_1,$ and d_1 are all nonzero positive real numbers or all nonzero negative real numbers. If $a_1, b_1, c_1, d_1, a_2, b_2, c_2,$ and d_2 are all nonzero positive real numbers, then the division of A_1 and A_2 is

$A_1 \oslash A_2 = (a_1, b_1, c_1, d_1; w_1) \oslash (a_2, b_2, c_2, d_2; w_2)$

$$= (\frac{a_1}{d_2}, \frac{b_1}{c_2}, \frac{c_1}{b_2}, \frac{d_1}{a_2}; \min(w_1, w_2)) \qquad 6.4$$

6.4.2 Fuzzy Weighted Average (FWA)

To consolidate the fuzzy values and fuzzy weights of all the important trust factors into a Fuzzy Weighted Average (FWA), (Schmucker, 1984) which will be used as an indication of the attractiveness of the situation. The higher the FWA value, the more trustworthy the agent is. Let R_j and W_j, where $j = 1, 2, ..., n$, respectively be the fuzzy rating and fuzzy weighting given to factor *j*. Then, the fuzzy weighted average is computed as

$$FWA = \sum_{j=1}^{n} (W_j \otimes R_j) / \sum_{j=1}^{n} W_j \qquad 6.5$$

6.4.3 Simple Center of Gravity Matching Method

Once the FWA has been computed, this value could be approximated by a similar close linguistic term from the Trust Value (TV, see Table 6.2). Several methods for matching the FWA with the corresponding TV have been proposed (Kosko, 1986; Rutter, 2000). The method of Simple Center of Gravity Method (SCGM) is based on the concept of medium curve. If A is a generalized trapezoidal fuzzy number, where $A = (a, b, c, d; w)$, then the value y^* of the COG point of A is

$$y_A^* = \begin{cases} \dfrac{w_A \times (\frac{c-b}{d-a} + 2)}{6}, & if \quad a \neq d \quad and \quad 0 < w_A \leq 1 \\ \dfrac{w_A}{2}, & if \quad a = d \quad and \quad 0 < w_A \leq 1 \end{cases} \qquad 6.6$$

If A is a generalized triangular fuzzy number, where $A = (a, b, c, d; w)$, then

$$y_A^* = \frac{w_A \times (\frac{b-b}{c-a} + 2)}{6}$$

$$= \frac{w_A \times (0+2)}{6} = \frac{w_A}{3}$$

The value of x^* of the COG point of A is

$$x_A^* = \frac{y_A^*(c+b) + (d+a)(w_A - y_A^*)}{2w_A} \qquad 6.7$$

The COG point of a generalized trapezoidal fuzzy number A is

$$COG(A) = (x_A^*, y_A^*). \qquad 6.8$$

6.4.4 Similarity Measure between Generalized Fuzzy Numbers

Assuming that there are two generalized trapezoidal fuzzy numbers A_1 and A_2 where $A_1 = (a_1, b_1, c_1, d_1; w_1)$, $A_2 = (a_2, b_2, c_2, d_2; w_2)$, $0 \le a_1 \le b_1 \le c_1 \le d_1 \le 1$, and $0 \le a_2 \le b_2 \le c_2 \le d_2 \le 1$.

First we obtain the COG points of A_1 and A_2. Then, the degree of similarity $S(A_1, A_2)$ could be calculated as follows:

$$S(A_1, A_2) = \left[1 - \frac{\sum_{i=1}^{4} |a_{1i} - a_{2i}|}{4}\right] \times (1 - |x_{A_1}^* - x_{A_2}^*|)^{B(S_{A_1}, S_{A_2})} \times \frac{\min(y_{A_1}^*, y_{A_2}^*)}{\max(y_{A_1}^*, y_{A_2}^*)} \qquad 6.9$$

where $S(A_1, A_2) \in [0,1]$, and $B(S_{A_1}, S_{A_2})$ is defined by

$$B(S_{A_1}, S_{A_2}) = \begin{cases} 1, & if \quad S_{A_1} + S_{A_2} > 0 \\ 0, & if \quad S_{A_1} + S_{A_2} = 0 \end{cases} \qquad 6.10$$

where S_{A_1} and S_{A_2} are the lengths of the bases of the generalized trapezoidal fuzzy numbers A_1 and A_2 respectively defined as follows:

$$S_{A_1} = d_1 - a_1$$
$$S_{A_2} = d_2 - a_2 \qquad 6.11$$

6.5 The Stag Hunt

In game theory, the stag hunt, also known as “trust dilemma," "assurance game," and "coordination game" represents a conflict between own safety and social cooperation. The name came from the work of the Swiss-born French philosopher Jean-Jacques Rousseau. In his work of 1755, *a Discourse on Inequality,* he described the situation where two men went hunting. Independently, each can choose to hunt a stag or hunt a hare. Without knowing the choice of the other, each one must choose his action. The two must cooperate with each other to hunt down a stag. Individually, each can hunt a hare by his own. The worth of a whole hare is less than the share of a stag. This has been considered as an important representation for social cooperation.

The odd thing about the stag hunt is that it shouldn't be a dilemma at all. One should certainly cooperate and hunt a stag. If both hunters do, both will get the best possible payoff. What complicate things is the possibility that one won't be so rational. If the one pursue a hare, the other wants to do too otherwise will end up with nothing.

In his writings, Rousseau, idealized ancient men and held civilization responsible for the majority of social problems. His philosophy in most is based on a rough and partially inaccurate conceptualization of prehistory. Part Two of *a Discourse on Inequality* states that the ancient human societies began when people forged temporary alliances for hunting.

If the issue is only about hunting a stag, then it is easy to realize that everyone should carry out his obligations. The issue arises when a hare happened to be within the reach of one of the hunters. The dilemma is now would he have gone off in pursuit of it, and therefore having caught his own prey, he wouldn’t care much about having caused his companion to lose his.

Among the most recent stag hunt dilemmas is the 1989 constitutional amendment to prohibit the burning of the U.S. flag and make it a federal crime. The main reason for objecting to this bill is that it violates of freedom of expression. On the other hand, the opposing party feared that if they voted against it and it passed, it would be used against them. Their opponents would show them "in favor of flag burning" and hence being unpatriotic in the next elections. According to Senator Joseph R. Biden, Jr., an opponent of the bill, "More than 45 senators would vote 'no' if they knew they were casting the deciding vote." Since the year 1995, the flag burning amendment has always gained enough votes to pass in the House of Representatives but not in the Senate. Fell four votes short of the required two-third majority, in 2000, the Senate voted 63–37 in favor of the amendment. Another attempt in 2006 fell one vote short.

The stag hunt could used to describe the ethical dilemma of the scientists who built the atomic bomb. In 1950, Harold Urey said of the hydrogen bomb: the world would be better off without the bomb. "I personally hope very much that the bombs will not explode, no matter how much effort is put into the project". Having no assurances that our enemies will not build it, we have to try to build it. It is better that we have the bomb rather than our enemies; better both sides have the bomb than just our enemies.

When recalling the head injury of professional hockey player Teddy Green in 1969, *Newsweek* stated: “Players will not adopt helmets by individual choice for several reasons. Chicago star Bobby Hull cites the simplest factor: "Vanity." But many players honestly believe that helmets will cut their efficiency and put them at a disadvantage, and others fear the ridicule of opponents. The use of helmets will spread only through fear caused by injuries like Green's – or through a rule making them mandatory.... One player summed up the feelings of many: "It's foolish not to wear a helmet. But I don't – because the other guys don't. I know that's silly, but most of the players feel the same way. If the league made us do it, though, we'd all wear them and nobody would mind."[1]

6.5.1 Formal Definition

In game theory terminology, the stag hunt is a game with two pure strategy Nash equilibria of which one is risk dominant and the other is payoff dominant. A payoff dominant equilibrium is a Nash equilibrium that pays at least as well as any other equilibrium for all players and for at least one player, it is strictly more. This is also referred to as pareto superior. This refinement of Nash equilibrium was introduced by Harsanyi and Selten in 1988 (Harsanyi and Selten, 1988). A risk dominant equilibrium is a player’s best response to a strategy profile of the other players or a probability distribution over these profiles. If each player assigns a uniform probability over the other’s pure strategies and s^* is the unique best response for both, then (s^*, s^*) is the risk dominant equilibrium (Harsanyi and Selten, 1988).

[1] Newsweek, October 6, 1969.

Table 6.3 shows the game in its generic form. In this table, while DM 1, or hunter 1, controls the row strategies, DM 2, or hunter 2, controls the column strategies. When each hunter selects a strategy, a state is formed, which is represented by a cell in the matrix. The double letters given at the top of each cell represent the strategies selection by each of the two DMs where the letters on the left and right stand for the strategies of DMs 1 and 2, respectively. For example, the cell given as SS which will be referred to by the encircled number 1, for simplicity sake, is the state in which both DM 1 and DM 2 pursue a stag. The two letters given in brackets in the bottom of a cell represent the preferences of DM 1 (left entry, uppercase letters) and DM 2 (right entry, lowercase letters), where a higher number means more preferred. Here it is assumed that $a > b \geq d > c$ and $A > B \geq D > C$. For instance, $A = a = 4$, $B = b = 3$, $C = c = 0$, and $D = d = 3$ (see Table 6.4). Games with a similar structure but without the risk dominant Nash equilibrium are still called stag hunt by some game theoretic. For example if $A = a = 2$, $B = b = 1$, $C = c = 0$, and $D = d = 1$. While the state HH remains a Nash equilibrium, it is no anymore a risk dominant. Nevertheless, many would still call this game a stag hunt.

Table 6.3: Generic form of the stag hunt.

		DM 2	
		Stag S	Hare H
DM 1	Stag S	① SS (A, a)	② SH (C, b)
	Hare H	③ HS (B, c)	④ HH (D, d)

Table 6.4: Example of the stag hunt.

		DM 2	
		Stag S	Hare H
DM 1	Stag S	① SS (4, 4)	② SH (0, 3)
	Hare H	③ HS (3, 0)	④ HH (3, 3)

While the prisoner's dilemma (refer to section 3.7 for more details) gained a lot of attention as the best game that represents the problem of social cooperation, some authors believe that the stag hunt represents an equally (or more) interesting context for studying social cooperation and its problems (Skyrms, 2003).

Due to the substantial relationship between the two games of prisoner's dilemma and stag hunt, many situations that have been described as prisoner's dilemma might also be interpreted as a stag hunt. In a similar fashion, some human interactions that seem like prisoner's dilemmas may in fact be a stag hunt. Assume that we have the prisoner's dilemma shown in Table 6.5. Players who defect when others tend to cooperate are usually punished for their defection. Imposing a -2 payoff as a punishment will turn this game of prisoner's dilemma into a stag hunt game.

Table 6.5: Prisoner's Dilemma.

		DM 2	
		Cooperate C	Defect D
DM 1	Cooperate C	① CC (4, 4)	② CD (0, 5)
	Defect D	③ DC (5, 0)	④ DD (3, 3)

6.5.2 Stag Hunt: Trust and Cooperation

There is a strong relationship between trust and cooperative actions and how one's attitudes and actions change when realizing the fact that one is being trusted or trusting someone else. Generally speaking, one could say that the presence of complete distrust will eliminate any chance for cooperation (Gambetta, 2000; Mares, 2001).

When trying to explain cooperative actions, it can't be studied in isolation of other influential factors like trust, importance of the situation, risk involved, and sharing some common goals. The presence of trust will eliminate or decrease fears of betrayal when engaging in a cooperative situation. Other factors like low risks and the importance of the situation could complement and support low levels of trust. On the other hand, having a high level of risk and having low importance when combined with low level of trust will reduce the chances of cooperation. One uses some mental shortcuts like trust to reduce the situation's complexity and the chances for risks thereafter (Luhmann, 1979).

As Rousseau (1955) stated: "If it was a matter of hunting a deer, everyone well realized that he must remain faithfully at his post; but if a hare happened to pass within the reach of one of them, we cannot doubt that he would have gone off in pursuit of it without scruple and, having caught his own prey, he would have cared very little about having caused his companions to lose theirs." Each needs assurances that he will not be betrayed by the other and ends up with nothing. To estimate the value of trust in this case, one uses the information in Tables 6.4 and 6.6. Table 6.6 shows the main criteria and sub-criteria for trust in the stag hunt game along with their corresponding fuzzy values and weights. Originally, most of these values are expressed in natural linguistic terms like saying the risk for this situation is high. From Table 6.2, one can see that the matching fuzzy interval for high is (0.8, 0.9, 1.0, 1.0; 1.0). Different decision makers would have different weights for the different factors. Using the information in Table 6.6 and employing Equation 6.5, one obtains the fuzzy values for the main criteria shown in Figure 6.5 as follows:

Information = (0.306, 0.522, 0.858, 1.308)

Circumstances = (0.312, 0.543, 0.909, 1.435)

Judgment = (0.236, 0.434, 0.765, 1.320)

Computing the Fuzzy Weighted Average using Equation 6.5, one obtains: FWA = (0.285, 0.555, 1.055, 1.935)

Table 6.6: Main and sub-criteria of trust and their fuzzy values and weights.

Criteria	Sub-Criteria	Fuzzy Values	Fuzzy Weights
Information			(0.7, 0.8, 0.9, 1.0; 1.0)
	Risk	(0.8, 0.9, 1.0, 1.0; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Uncertainty	(0.6, 0.7, 0.8, 0.9; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Dependence	(0.8, 0.9, 1.0, 1.0; 1.0)	(0.5, 0.6, 0.7, 0.8; 1.0)
	Reputation	(0.4, 0.5, 0.6, 0.7; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
Circumstances			(0.7, 0.8, 0.9, 1.0; 1.0)
	Importance	(0.6, 0.7, 0.8, 0.9; 1.0)	(0.5, 0.6, 0.7, 0.8; 1.0)
	Value	(0.8, 0.9, 1.0, 1.0; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Alternatives	(0.6, 0.7, 0.8, 0.9; 1.0)	(0.5, 0.6, 0.7, 0.8; 1.0)
	Deadlines		
Judgment			(0.7, 0.8, 0.9, 1.0; 1.0)
	Attractiveness	(0.6, 0.7, 0.8, 0.9; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Actions & Statments		
	Feelings & Emotions	(0.4, 0.5, 0.6, 0.7; 1.0)	(0.3, 0.4, 0.5, 0.6; 1.0)

Using Equations 6.6, 6.7, 6.8, and 6.11, the COG and base points for the five-member fuzzy linguistic set in Table 6.2 are shown in Table 6.7.

Table 6.7: COG and base points for fuzzy trust values.

Linguistic term	COG points	Base
Low	(0.15,0.389)	0.30
Fairly low	(0.35,0.389)	0.30
Medium	(0.55,0.389)	0.30
Fairly high	(0.75,0.389)	0.30
High	(0.85,0.500)	0.20

In a similar fashion, the COG (FWA) = (0.993, 0.383) and the base S= 1.65. According to Equation 6.10, one can see that values of $B(S_R, S_{low})$, $B(S_R, S_{fairly-low})$, $B(S_R, S_{medium})$, $B(S_R, S_{fairly-high})$, and $B(S_R, S_{high})$ are all equal to 1. Based on Equation 6.9, the degree of similarity between the Fuzzy Weighted Average (FWA) and the five-member linguistic terms are:

$S(R, low) = 0.030$.

$S(R, fairly-low) = 0.138$.

$S(R, medium) = 0.294$.

$S(R, fairly-high) = 0.420$.

$S(R, high) = 0.370$.

Because $S(R, fairly-high)$ has the highest value (0.420), Fuzzy Weighted Average (FWA) is translated into a fairly high value. In this case, we would say that the degree of similarity is 0.420. In other words, the trust index (based on the available information) is fairly high.

6.5.3 Stability Analysis

Stag Hunt can be solved using different solution concepts like Nash stability, general metarationality (GMR), symmetric metarationality (SMR), and sequential stability (SEQ) (see Fang et al., 1997 for precise definitions of these solution concepts along with original references). The results obtained using these different solution concepts are provided in Table 6.8. In this table, a “U” stands for a state that is unstable for a particular player. In other words, a player could move unilaterally to another state that will produce a better payoff without a credible sanction being levied by the opponent player. An “S” stands for a state that is stable for a particular player. It could be stable because there are no unilateral improvements by that particular player or for any unilateral improvement the opponent player can invoke a sanction according to the way a sanction is defined for a given solution concept. A state that is stable for both players constitutes an equilibrium state denoted by E.

According to Nash stability, the equilibrium point will be SS and HH. A Nash equilibrium is the situation where it is not advantageous for either player to move to another state unilaterally because it will produce a worse payoff or at least not better than the current one. Equilibrium SS represents the case were both hunters will cooperate and together pursue a stag. However, there is always a chance that one of them will go after a hare and securing his gain while letting his companion loss his. Whether it is mentioned explicitly or not, only the presence of some trust will eliminate this fear of betrayal. Howard (1971), derived the solution concepts of GMR and SMR which realistically predict state both SS and HH as equilibria. It works as corrective action in case of defect from the other. It gives the player the chance to counteract the action of the opponent. The SEQ solution concept developed by Fraser and Hipel (1984) also forecasts SS, along with HH as equilibria. An attractive feature of SEQ stability is that only credible sanctioning is permitted by the sanctioning player. As an example of how to calculate SEQ stability, consider state HS from player 1's viewpoint. As shown in Table 5.2, if the game where at state HS, player 1 can unilaterally improve from state HS to state SS by changing his strategy from hare to stag (notice that the ordinal payoff for player 1 is 3 in state HS versus 4 at state SS). However, player 2 has no unilateral improvement from state SS. Since there is no credible sanction for this move by player 2 and hence the new state SS is more preferred by both players, state HS is not SEQ stable for player 1. Since both SS and HH have no UIs for both players, they both constitute SEQ equilibria.

If one considers that in some real life situations a move may be irreversible and can't be taken back once it is invoked, a state may no longer be stable according to a particular solution concept. In fact, the graph model for conflict resolution can directly account for irreversible moves (Fang et al., 1993). For example, once a hunter pursued a hare, he missed the chance to help his companion hunt down a stag. Since each has an incentive to secure a hare rather than waiting to hunt a stag, in the absence of trust, each will fear that the other will betray him. In this case, each will behave in a rational manner and defect. The outcome in this case is that both will have a hare and therefore a less payoff when compared to a share of stag. The absence of trust weakens the coalition between the two hunters though it is clear that it is better for both if they cooperate.

On the other hand, if trust is present, it eliminates any fears of betrayal even if they don't know each other preferences. It acts like a binding agreement that strengthens and supports the coalition between the two. The stronger is the trust, the stronger is the coalition. Each will sacrifice a chance for a hunting a hare (secured gain) for a more rewarding share for both. Any short term benefits gained from betraying the other will result in a long term loss in any future interaction. In such repeated encounters, building a good reputation is crucial for future long lasting benefits.

Table 6.8: Stag hunt stability analysis using different solution concepts.

Solution Concept	State	Player 1	Player 2	Equilibria
Nash	SS	S	S	E
	SH	U	U	
	HS	U	U	
	HH	S	S	E
GMR	SS	S	S	E
	SH	U	U	
	HS	U	U	
	HH	S	S	E
SMR	SS	S	S	E
	SH	U	U	
	HS	U	U	
	HH	S	S	E
SEQ	SS	S	S	E
	SH	U	U	
	HS	U	U	
	HH	S	S	E

6.6 Groundwater Contamination Conflict

Recalling the groundwater contamination conflict presented and solved in section 4.2, the equiliberia states are 1, 4, 5, 8, and 9. While state 1 represents the status quo and state 9 represents the worst case scenario, the remaining states (4, 5, and 8) lies somewhere in between. While some of the equilibrium states are more preferred than others, the possibility of one of them occurring depends on the strategy choices by some or all the players involved. It might be in the interest of some of the players to cooperate and together move to a more preferred equilibrium state given that this move is not sanctioned by a move levied by a non-coalition member. This may require some of the coalition members to move to an intermediate state that might not be preferred when compared to the current state by some of the players. This entitles some risks and fear of betrayal.

States 5 and 8 are more preferred by all players to state 9. State 8 is more preferred by both MoE and UR to state 5. It is not possible for any of them to move from state 5 to state 8 on an individual basis. In order to do so, UR needs to move first to state 7, which is less preferred by UR, and, subsequently, MoE can move from state 7 to state 8. In case MoE doesn't move to state 8, UR will end up in a less preferred situation and might need to abandon its entire operations in Elmira.

Table 6.9 shows the evolution of the Elmira conflict from the status quo, to a transitional non-cooperative equilibrium and ultimately to a final cooperative coalition equilibrium. The status quo is state number 1 which is the most preferred state for UR but less preferred by both MoE and LG. As indicated by the arrow connecting states 1 and 5 in Table 6.9, LG can unilaterally cause the conflict to move from state 1 to 5 by changing its strategy from not insisting that the original be adopted to insisting that it be applied. In fact, because LG prefers state 5 to 1 (see the third row in Table 4.3 in section 4.2), this change in state constitutes a unilateral improvement for LG. Though state number 5 is not the most preferred for both MoE and LG, it is still more preferred than state number 1. From state number 5, it is not possible for any player to improve unilaterally. In order to move from state 5 to state 8, two players need to move together (MoE and UR). This requires moving to an intermediate state; either 6 or 7. If MoE initiates the move, then the intermediate state is 6. Since state 6 is less preferred by UR compared to state 8, it is natural that UR wants to move jointly with MoE to directly reach state 8. If UR initiates the move from state 5, then the intermediate state is state 7 which is the most preferred state for MoE. Accordingly, for the coalition consisting of MoE and UR to jointly improve from state 5 to 8, the decision makers must trust one another.

As explained above, if UR moves from state 5 without the participation of MoE, the less preferred state 7 will be formed, which happens to be MoE's most preferred state. Hence, MoE has an incentive to remain at state 7 if it fools UR into trustingly moving to state 7. If it does not wish to abuse the trust required of UR, it will remain in the coalition and select its strategy such that state 8 is reached. Of course, both MoE and UR prefers state 8 to state 5 and this joint gain helps solidify trust and vice versa.

Table 6.9: Evolution of the Elmira Conflict.

	Status Quo	Transitional Non-cooperative Equilibrium	Cooperative Equilibrium
MoE: 1. Modify	N	N →	Y
UR: 2. Delay	Y	Y →	N
3. Accept	N	N →	Y
4. Abandon	N	N	N
LG: 5. Insist	N →	Y	Y
State Number	1	5	8

In the foregoing situation, both MoE and UR have an incentive not to carry out their obligations and move the coalition from the intermediate state 5 to the final state. Only if there is a minimum level of trust, will each execute their obligations and sacrifice individual benefits for a more preferred one by both (but might be less preferred for one when compared to the intermediate state). UR needs to calculate the trust index for both MoE and LG to ascertain if it worth forming a coalition with any of them. To calculate the trust index between UR and MoE, using the information in Table 6.10 and employing Equation 6.5, one obtains the fuzzy values for main criteria as follows:

Information = (0.371, 0.568, 0.845, 1.240)

Circumstances = (0.460, 0.650, 0.862, 1.130)

Judgment = (0.560, 0.800, 1.125, 1.428)

Table 6.10: Trust index for UR with MoE.

Criteria	Sub-Criteria	Fuzzy Values	Fuzzy Weights
Information			(0.9, 1.0, 1.0, 1.0; 1.0)
	Risk	(0.6, 0.7, 0.8, 0.9; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Uncertainty	(0.6, 0.7, 0.8, 0.9; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Dependence	(0.4, 0.5, 0.6, 0.7; 1.0)	(0.5, 0.6, 0.7, 0.8; 1.0)
	Reputation	0	0
Circumstances			(0.7, 0.8, 0.9, 1.0; 1.0)
	Importance	(0.6, 0.7, 0.8, 0.9; 1.0)	(0.9, 1.0, 1.0, 1.0; 1.0)
	Value	(0.4, 0.5, 0.6, 0.7 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Alternatives	(0.8, 0.9, 1.0, 1.0; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Deadlines	0	0
Judgment			(0.5, 0.6, 0.7, 0.8; 1.0)
	Attractiveness	(0.8, 0.9, 1.0, 1.0; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Actions & Statements		
	Feelings & Emotions	0	0

Computing the Fuzzy Weighted Average using Equation 6.5, one obtains: FWA = (0.325, 0.598, 1.062, 1.810). Employing equations 6.6, 6.7, 6.8, and 6.11, the COG (FWA) = (0.976, .038) and the base S = 1.48. According to Equation 6.10, one can see that values of $B(S_R, S_{low})$, $B(S_R, S_{fairly-low})$, $B(S_R, S_{medium})$, $B(S_R, S_{fairly-high})$, and $B(S_R, S_{high})$ are all equal to 1. Based on Equation 6.9, the degree of similarity between the Fuzzy Weighted Average (FWA) and the five-member linguistic terms are:

$S(R, low) = 0.035.$

$S(R, fairly-low) = 0.149.$

$S(R, medium) = 0.321.$

$S(R, fairly-high)$ = 0.470.

$S(R, high)$ = 0.413.

From these calculations, one concludes that the trust index between UR and MoE is Fairly High.

To calculate the trust index between UR and LG, using the information in Table 6.11 and employing Equation 6.5, one obtains the fuzzy values for main criteria as follows:

Information = (0.307, 0.488, 0.745, 1.120)

Circumstances = (0.260, 0.407, 0.577, 0.826)

Judgment = (0.140, 0.267, 0.450, 0.714)

Computing the Fuzzy Weighted Average using Equation 6.5, one obtain: FWA = (0.165, 0.344, 0.665, 1.265). Employing equations 6.6, 6.7, 6.8, and 6.11, the COG (FWA) = (0.635, .382) and the base S = 1.10. According to Equation 6.10, one can see that values of $B(S_R, S_{low})$, $B(S_R, S_{fairly-low})$, $B(S_R, S_{medium})$, $B(S_R, S_{fairly-high})$, and $B(S_R, S_{high})$ are all equal to 1. Based on Equation 6.9, the degree of similarity between the Fuzzy Weighted Average (FWA) and the five-member linguistic terms are:

$S(R, low)$ = 0.273.

$S(R, fairly-low)$ = 0.508.

$S(R, medium)$ = 0.670.

$S(R, fairly-high)$ = 0.588.

$S(R, high)$ = 0.376.

From these calculations, one concludes that the trust index between UR and LG is Medium.

Table 6.11: Trust index for UR with LG.

Criteria	Sub-Criteria	Fuzzy Values	Fuzzy Weights
Information			(0.9, 1.0, 1.0, 1.0; 1.0)
	Risk	(0.4, 0.5, 0.6, 0.7; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Uncertainty	(0.4, 0.5, 0.6, 0.7; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Dependence	(0.6, 0.7, 0.8, 0.9; 1.0)	(0.5, 0.6, 0.7, 0.8; 1.0)
	Reputation	0	0
Circumstances			(0.7, 0.8, 0.9, 1.0; 1.0)
	Importance	(0.4, 0.5, 0.6, 0.7; 1.0)	(0.9, 1.0, 1.0, 1.0; 1.0)
	Value	(0.4, 0.5, 0.6, 0.7 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Alternatives	(0.2, 0.3, 0.4, 0.5; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Deadlines	0	0
Judgment			(0.5, 0.6, 0.7, 0.8; 1.0)
	Attractiveness	(0.2, 0.3, 0.4, 0.5; 1.0)	(0.7, 0.8, 0.9, 1.0; 1.0)
	Actions & Statements	0	0
	Feelings & Emotions	0	0

Given the preferences for each decision maker and by comparing the two trust indices, one can envision UR trusting MoE and, hence, forming a coalition with MoE is more promising than forming a one with the LG.

In reality, UR formed a coalition with MoE and accepted the Control Order and at the same time the MoE modified it to be more favorable to UR though such an action is not accepted by the local government. This coalition between MoE and UR caused the game to move from state 5 to state 8 (Kilgour et al., 2001).

6.7 Summary

The foregoing analysis for the stag hunt and the groundwater contamination conflict show the effectiveness of the fuzzy approach in modeling trust characteristics of DMs that other approaches fail to accomplish on their own. More specifically, the fuzzy approach to trust complements and strengthens the arguments that solution concepts may suggest as to when it is in DMs' interests to form coalitions in order to benefit coalition members. For example, even though all coalition members may fare better within a coalition, one or more coalition members may still be tempted to act independently because they think they may gain even more if they behave selfishly. The famous stag hunt dilemma and the realworld environmental conflict were employed to illustrate and explain how trust can provide useful insights about human behavior under conflict.

The fuzzy approach to trust adds a more realistic dimension to the study of conflict by capturing how the DMs would think and behave in a situation in which a decision to cooperate is to be made. It also shows the strong relationship between trust and cooperative actions and how one's strategy choices change when realizing the fact that one is being trusted or trusting the other party. Generally speaking, one could say that the presence of complete distrust will eliminate any chance for cooperation. In the Elmira Conflict, the idea of coalition analysis gives a partial answer to the question of whether to cooperate or not. While the coalition analysis highlights the possible improvements by forming a coalition, it doesn't give a full answer on how and why the coalition will form nor which coalitions take place. Through the use of the fuzzy approach and the introduction of the concept of trust among coalition members, this research provides an answer to these questions. Chapter 7 concludes by summarizing the contribution of this thesis and the future work.

Chapter 7

Conclusions and Future Research

The introduction of the new fuzzy preference representation extends the applicability of the graph model for conflict resolution to areas that were not previously attainable. The introduction of the trust concepts and modeling could help understand, assess, and evaluate the feasibility of forming a coalition in order to improve to a more preferred equilibria that couldn't be reached unilaterally by a single DM.

The Graph Model for Conflict Resolution is an intuitive, flexible, and at the same time efficient paradigm for modeling and analyzing strategic conflict (Kilgour et al., 2001; Fang et al., 1993). However, until now it didn't take into consideration preference uncertainties. Additionally, when the trust research in this thesis is utilized, the graph model methodology is now able to explain why coalitions form and which ones are most likely to take place in practice.

7.1 Summary of Contributions

In real life, the actions and choices by others are anything but certain. One expects people to act or behave in a certain way but, in reality they are free agents who may not behave according to others' expectations. Incorporating the concept of uncertainty into the Graph Model for Conflict Resolution will help account for those uncertainties. The key contributions of this research are to introduce a fuzzy modeling of preference uncertainties and trust within the graph model and show how the presence of trust affects DMs' choices in cooperative games. Those key research contributions are shown in Figure 7.1.

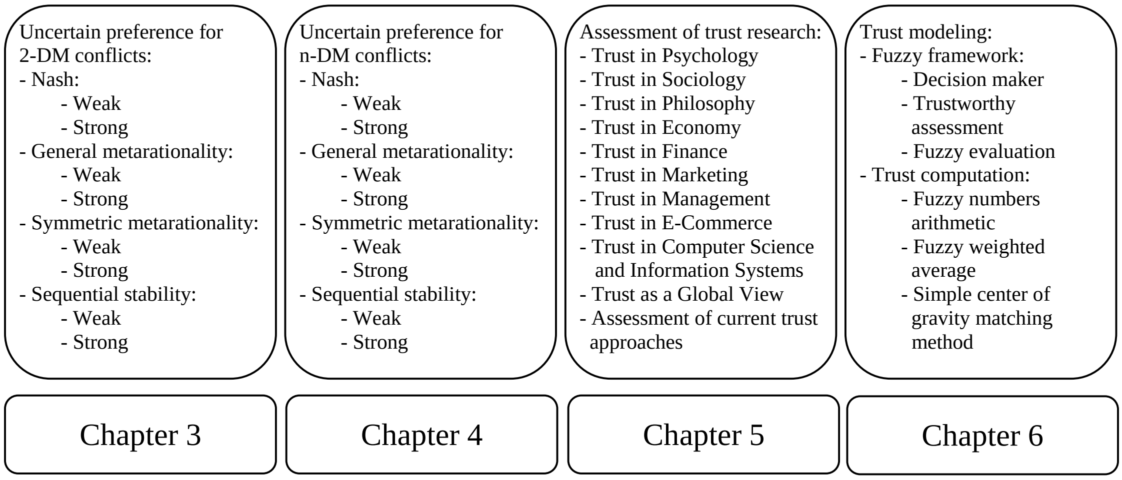

Figure 7.1: Major research contributions of the thesis.

In conflicts between just two participants, one only needs to worry about possible countermoves by the opponent depending on the solution concept under investigation. Chapter three introduces the concept of fuzzy preferences. The concepts of unilateral moves and unilateral improvements are modified in accordance with the new fuzzy preferences. The four solution concepts of Nash, general metarationality, symmetric metarationality, and sequential stability are redefined to accommodate fuzziness in preferences. One important distinction between the newly defined solution concepts and the regular ones is that the strength of Nash stability, for which there are no UIs, is defined based on the strength of the preference. The interrelationships among the different concepts are highlighted in Section 3.4. In Section 3.5, the concept of the fuzzy binary relationship is introduced and the relationship between the fuzzy and non-fuzzy preferences is examined afterwards. To show the applicability of the newly proposed fuzzy preference representation, the well-known game of Prisoner's Dilemma is examined from a fuzzy perspective for both the ordinal and cardinal cases.

The newly introduced solution concepts in Chapter 3 are generalized for when there are more than two DMs in Chapter 4. This requires examining all possible moves and countermoves by all the sanctioning DMs which calls for defining them as a sequence of allowable moves rather than a single one. The groundwater contamination conflict is employed to show how the new solution concepts for multiple participants are applied in practice.

Upon reaching euqilibria or a conflict resolution, it might be in the best interests of some of the DMs to form a coalition and move together to a better cooperative equilibrium which is not attainable on an individual basis. However, this may involve some risks as the new coalition might require some of its members to move to an intermediate state that might be less preferred by some members while more preferred by others when compared to the original or the final states. Only the presence of some minimum level of trust will eliminate any fears of betrayal and ensure all the coalition members that each and every one will fulfill his or her obligations towards reaching the final goal of the coalition.

To enhance the understanding of the concept of trust and help to better model it, Chapter 5 starts with a thorough study and assessment of trust among different disciplines. This study explores how the different disciplines envision trust differently and why. Current issues with the current methods for trust assessment and modeling have been highlighted. This research calls for developing a general data test set besides agreeing on a general testing framework for trust.

Chapter 6 proposes a fuzzy structure for modeling trust in order to capture as much as possible the trust dynamics and multidimensionality. It is simple, intuitive, and uses linguistic terms that are easy to express and understand. This enables humans and agents involved in conflicts or cooperative actions to evaluate the trustworthiness of others before engaging with them in an interaction. As a representative generic game for so many real life encounters, the game of stag hunt was employed to demonstrate the applicability of the proposed modeling for two-DM disputes. Between the fear of the worst and the hope for the best, the aquifer water contamination conflict represents an ideal situation for applying the trust methodology for conflicts having more than 2 DMs.

7.2 Future Research

Despite the theoretical advances of introducing the concept of the preference uncertainties in Graph Model for Conflict Resolution for both 2 and n-DM cases, there are some promising issues that need to be investigated further:

- The current version of the decision support system GMCR II doesn't take into consideration any preference uncertainties. Implementing the new fuzzy representation within the framework of the Graph Model for Conflict Resolution will extend the applicability of the system.
- It is worth investigating the applicability of the new fuzzy preference representation in other solution concepts such as limited-move stability.

- Currently, it is assumed that all factors shown in Figure 6.1 contribute equally to trust. In fact, this might not be true. Some factors, which are considered important issues to trust, may in fact contribute less to trust.
- Another issue is the interaction among the different factors. Some factors might be influential to others but do not directly affect trust while others might have a direct impact on the value of trust. A well-designed survey or empirical study and the employment of some stochastic analysis, multivariate inference, or hybrid fuzzy stochastic models are needed to further highlight relationships among the different factors and show which ones contribute more to trust, and whether or not this is done directly or indirectly.

Bibliography

Abdul-Rahman, A., and Hailes, S. (2000). "Supporting trust in virtual communities," *Proceedings of the Hawaii International Conference on System Sciences*, Hawaii :Maui.

Abdul-Rahman, A. and Hailes, S. (1997a). "A distributed trust model," *Proceedings of the New Security Paradigms Workshop*.

Abdul-Rahman, A. and Hailes, S. (1997b). "Using recommendations for managing trust in distributed systems," *Proceedings of IEEE International Conference on Communication*.

Aberer, K., and Despotovic, Z. (2001). "Managing trust in a peer-2-peer information system," *Proceedings of the Tenth International Conference on Information and Knowledge Management (CIKM01)*.

Abrams, M. (1995). "Trusted system concepts," *Computers and Security* 14(1), 45– 56.

Acrement, B. (2002). "Elements for building trust: Do your management skills measure up?." Retrieved November 12, 2006, from http://www.imakenews.com/smei/e_article000051474.cfm. Checked on January 23, 2006.

Adar, E. and Huberman, B. (2000, October). "Free riding on Gnutella." Retrieved November 15, 2006, from http://www.firstmonday.org/issues/issue5_10/adar/. Checked on January 23, 2007.

Advogato (2000). "Advogato's trust metric." Retrieved Novemeber 15, 2006, from http://www.advogato.org/trust-metric.html. Checked on January 23, 2007.

AL-Mutairi, M., Hipel, K., and Kamel, M. (2005a). "Cooperation from a Fuzzy Perspective," *Proceedings 17th World Congress Scientific Computation, Applied Mathematics and Simulation (IMACS2005),* France: Paris.

AL-Mutairi, M., Hipel, K., and Kamel, M. (2005b). "Modeling Trust Using a Fuzzy Logic Approach," *Proceedings of International Symposium on Systems and Human Science*, USA: San Francisco.

AL-Mutairi, M., Hipel, K., and Kamel, M. (2006a). "Conflict Resolution with Fuzzy Preferences," *Proceedings International Symposium on Systems and Human Science*, Austria: Vienna.

AL-Mutairi, M., Hipel, K., and Kamel, M. (2006b). "Fuzzy Preferences in Conflicts among Multiple Participants," *Fuzzy Sets and Systems*, forthcoming.

AL-Mutairi, M., Hipel, K., and Kamel, M. (2006c). "Trust and Cooperation from a Fuzzy Perspective," *Mathematics and Computers in Simulation*, forthcoming.

AL-Mutairi, M., Hipel, K., and Kamel, M. (2006d). "Two-Decision-Maker Conflict Resolution with Fuzzy Preferences," *Reliability Engineering and System Safety*, forthcoming.

AL-Mutairi, M., Hipel, K., and Kamel, M. (2007). "Trust: Different Views, One Goal," *Business and Technology Journal*, forthcoming.

Amazon (2002). "Amazon Auctions." Retrieved November 15, 2006, from http://auctions.amazon.com. Checked on January 23, 2007.

Ashri, R., Ramchurn, S., Sabater, J., Luck, M., and Jennings, N. (2005). "Trust Evaluation Through Relationship Analysis," *Proceedings of the 4th International Joint Conference on autonomous Agents and Multi-Agent Systems*.

Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books.

Axelrod, R., and Hamilton, W. (1981). "The Evolution of Cooperation," *Science* 211, 1390-1396.

Bacharach, M., and Gambetta, D. (2001). "Trust in Society," in K. Cook (Ed.), *Trust in signs*. New York: Russell Sage Foundation.

Baier, A. (1986). "Trust and Antitrust," *Ethics* 96, 231-260.

Barber, B. (1983). *Logic and Limits of Trust*. Rutgers University Press.

Barber, K., and Kim, J. (2001). "Belief Revision Process based on Trust: Simulation Experiment," *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies*, Canada: Montreal.

Barrett, C., Pattanaik, P., and Salles, M. (1990). "On Choosing Rationally when Preferences are Fuzzy," *Fuzzy Sets Systems*, 34, 197-212.

Becker, M., and Sewell, P. (2004). "Cassandra: Distributed Access Control Policies with Tunable Expressiveness," *Proceedings of the 5th IEEE International Workshop on Policies for Distributed Systems and Networks*.

Ben-Haim, Y. (2001). *Information Gap Decision Theory, Decisions under Severe Uncertainty*. Academic Press.

Ben-Haim, Y., and Hipel, K. (2002). "The Graph Model for Conflict Resolution with Information-Gap Uncertainty in Preferences," *Applied Mathematics and Computation* 126(2-3), 319-340.

Bendor, J. (1993). "Uncertainty and the Evolution of Cooperation," *Journal of Conflict Resolution* 37, 709-733.

Bendor, J., Kramer, R., and Swistak, P. (1996). "Cooperation Under Uncertainty: What is New, What is True and What is Important?," *American Sociological Review* 61, 333-338.

Bennett, P. (1977). "Toward a Theory of Hypergames," *Omega* 5(6), 749-751.

Bennett, P. (1980). "Hypergames: Development a Model of Conflicts," *Futures* 12(6), 489-507.

Bennett, P. (1995). "Modeling Decision in International Relations: Game Theory and Beyond," *Mershon International Studies Review*, 39, 19–52.

Bennett, P. (1998). "Confrontation Analysis as a Diagnostic Tool," *European Journal of Operational Research*, 109, 465-482.

Bennett, P., and Howard, N. (1996). "Rationality, Emotion, and Preference Change: Drama-Theoretic Models of Choice," *European Journal of Operational Research* 92, 603-614.

Berners-Lee, T. (1999). *Weaving the Web*. Harper.

Beth, T., Borcherding, M., and Klein, B. (1994). "Valuation of trust in open networks," *Proceedings of the 3rd European Symposium on Research in Computer Security*.

Bhattacherjee, A. (2002). "Individual Trust in Online Firms: Scale Development and Initial Trust," *Journal of Management Information Systems* 19(1), 213–243.

Binmore, K. (1992). *Fun and Games: a Text on Game Theory*. Lexington, D.C. Health.

Binmore, K., and Samuelson, L. (1992). "Evolutionary Stability in Repeated Games Played by Finite Automata," *Journal of Economic Theory* 57, 278-305.

Bizer, C., Cyganiak, R., Gauss, T., and Maresch, O. (2005). “The TriQL.P Browser: Filtering Information Using Context-Content and Rating-Based Trust Policies,” *Proceedings of the 4th International Semantic Web Conference*.

Blaze, M., Feigenbaum, J., Ioannidis, J., and Keromytis, A. (1999). “The Role of Trust Management in Distributed System Security,” *Lecture Notes in Computer Science* 1603, 185–210.

Blaze, M., Feigenbaum, J., and Lacy, J. (1996). “Decentralized Trust Management,” *Proceedings of IEEE Symposium on Security and Privacy*.

Blomqvist, K. (1997). “The Many Faces of Trust,” *Scandinavian Journal of Management*, 13, 271-286.

Boerlijst, M., Nowak, M., and Sigmund, K. (1997). “The Logic of Contrition,” *Journal of Theoretical Biology* 185, 281-293.

Boeyen, S., Ellison, G., Karhuluoma, N., MacGregor, W., Madsen, P., Sengodan, S., Shinkar, S., Thompson, P., and Linn, J. (2003). “Liberty Trust Models Guidelines,” in: J. Linn (Ed.), *Liberty Alliance Project*. Liberty Alliance, Draft Version 1.0–15 Edition.

Boyd, R. (1989). “Mistakes Allow Evolutionary Stability in the Repeated Prisoner's Dilemma Game,” *Journal of Theoretical Biology* 136, 47-56.

Boyd, R., and Lorberbaum, J. (1987). “No Pure Strategy is Evolutionarily Stable in the Repeated Prisoner's Dilemma Game,” *Nature* 327, 58-59.

Bower, A., Garber, S., and Watson, J. (1996). “Learning About a Population of Agents and the Evolution of Trust and Cooperation,” *International Journal of Industrial Organization*15, 165-190.

Brainov, S. and Sandholm, T. (2002). “Contracting With Uncertain Level of Trust,” *Computational Intelligence* 18, 501-514.

Brams, S. (1994). *Theory of Moves*. Cambridge: Cambridge University Press.

Brams, S., and Wittman D. (1981). “Nonmyopic Equilibria in 2 x 2 Games,” *Conflict Management and Peace Science* 6(1), 39-42.

Braynov, S., and Sandholm, T. (2002). “Incentive Compatible Mechanism for Trust Revelation,” *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

Breban, S. and Vassileva, J. (2002). “Using Inter-Agent Trust Relationships for Efficient Coalition Formation,” *Proceedings of the 15th Conference of the Canadian Society for Computational Studies of Intelligence on Advances in Artificial Intelligence*.

Brien, A. (1998). “Professional Ethics and the Culture of Trust,” *Ethics and Information Technology* 17(4), 391-409.

Bromley, D. (1993). *Reputation, Image and Impression Management*. New York: JohnWiley&Sons.

Bryant, J. (2003). *The Six Dilemmas of Collaboration: Inter-organisational Relationships as Drama*. Chichester, England: John Wiley.

Brainov, S., and Sandholm, T. (1999). “Contracting with Uncertain Level of Trust,” *EC ’99: Proceedings of the 1st ACM conference on Electronic commerce*.

Brin, S., and Page, L. (1998). “The Anatomy of a Large-Scale Hypertextual Web Search Engine,” *Computer Networks and ISDN Systems* 30(1–7), 107–117.

Buchegger, S., and Boudec, J.-Y. Le (2003). “A Robust Reputation System for Mobile Ad-hoc Networks,” *Technical Report IC/2003/50*, EPFL-IC-LCA.

Buchegger, S., and Boudec, J.-Y. Le (2003). “The Effect of Rumor Spreading in Reputation Systems for Mobile Ad-hoc Networks,” *Proceedings of the Workshop on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*.

Buckley, J. (1984). “The Multiple Judge, Multiple Criteria Ranking Problem: a Fuzzy Set Approach,” *Fuzzy Sets and Systems* 13, 25-37.

Buskens, V. (1998). “The Social Structure of Trust,” *Social Networks* 20, 265–289.

Cahill, V., Shand, B., Gray, E., Dimmock, N., Twigg, A., Bacon, J., English, C., Wagealla, W., Terzis, S., Nixon, P., Bryce, C., Serugendo, G., Seigneur, J., Carbone, M., Krukow, K., Jensen, C., Chen, Y., and Nielsen, M. (2003). “Using trust for secure collaboration in uncertain environments,” *Pervasive Computing* 2(3), 52– 61.

Carbo, J., Molina, J., and Davila, J. (2002). “Comparing Predictions of SPORAS vs. a Fuzzy Reputation Agent System,” *Third International Conference on Fuzzy Sets and Fuzzy Systems*, Switzerland: Interlaken.

Carbone, M., Nielsen, M., and Sassone, V. (2003). “A formal model for trust in dynamic networks,” *Proceedings of International Conference on Software Engineering and Formal Methods (SEFM’03)*, Australia: Brisbane.

Carter, J., Bitting, E., and Ghorbani, A. (2002). “Reputation Formalization for an Information-Sharing Multi-Agent Sytem,” *Computational Intelligence* 18(2): 515–534.

Castelfranchi, C., and Falcone, R. (1998). “Principles of Trust for MAS: Cognitive Anatomy, Social Importance, and Quantification,” *Proceedings of the International Conference on Multi-Agent Systems (ICMAS’98)*, France: Paris.

Castelfranchi, C., and Tan, Y. (2001). *Trust and Deception in Virtual Societies*. Kluwer Academic Publishers.

Castelfranchi, C., Falcone, R., and Pezzulo, G. (2003). “Trust in Information Sources as a Source for Trust: a Fuzzy Approach,” *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, USA: New York.

Celentani, M., Fudenberg, D., Levine, D., and Psendorfer, W. (1966). “Maintaining a Reputation Against a Long-Lived Opponent,” *Econometrica* 64(3), 691–704.

Chang, R. (1998). “Credible Monetary Policy with Long-Lived Agents: Recursive Approaches,” *Journal of Economic Theory* 81, 431-461.

Chari, V., and Kehoe, P. (1990). “Sustainable Plans,” *Journal of Political Economy* 98, 783–802.

Chen, G., and Pham, T. (2001). *Introduction to Fuzzy Sets, Fuzzy Logic, and Fuzzy Control Systems*. RCR Press.

Chen, M., and Singh, J. (2001). “Computing and Using Reputations for Internet Ratings,” *Proceedings of the Third ACM Conference on Electronic Commerce (EC’01)*.

Chen, S., and Chen, S. (2003). “Fuzzy Risk Analysis Based on Similarity Measures of Generalized Fuzzy Numbers,” *IEEE Transactions on Fuzzy Systems* 11(1), 45-56.

Chen, S., and Hwang, C. (1992). *Fuzzy Multiple Attribute Decision Making Methods and Applications*. New York: Springer.

Chirita, P., Nejdl, W., Schlosser, M., and Scurtu, O. (2004). “Personalized Reputation Management in P2P Networks,” *Proceedings of the Trust, Security and Reputation Workshop held at the 3rd International Semantic Web Conference*.

Chklovski, T., Gil, Y., Ratnakar, V., and Lee, J. (2003). "Trellis: Supporting Decision Making via Argumentation in the Semantic Web," *Proceedings of the 2nd International Semantic Web Conference*.

Chopra, K. and Wallace, W. (2002). "Trust in Electronic Environments," *Proceedings of the 36th Hawaii International Conference on System Sciences*, Hawaii.

Chu, Y., Feigenbaum, J., LaMacchia, B., Resnick, P., and Strauss, M. (1997). "Referee: Trust Management for Web Applications," *World Wide Web Journal* 2(3), 127-139.

Ciolek, M. (1996). "The Six Quests for the Electronic Grail: Current Approaches to Information Quality in WWW Resources." Retrieved October 30, 2004, from http://www.ciolek.com/PAPERS/six-quests1996.html. Checked on Jaunary 23, 2007.

Clarke, C., Cormack, G., and Lynam, T. (2001). "Exploiting Redundancy in Question Answering," *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*, USA: New York.

Clausen, A. (2004). "The cost of attack of PageRank," *Proceedings of The International Conference on Agents, Web Technologies and Internet Commerce (IAWTIC'2004)*, Australia: Gold Coast.

Conte, R., and Paolucci, M. (2002). *Reputation in Artificial Societies: Social Beliefs for Social Order*. Kluwer Academic Publishers.

Cornelli, F., Damiani, E., Di Vimercati, D., Paraboschi, S., and Samarati, P. (2002). "Choosing Reputable Servents in a P2P Network," *Proceedings of the eleventh international conference on World Wide Web (WWW'02)*, USA: New York.

Corritore, C., Wiedenbeck, S., and Kracher, B. (2001). "The Elements of Online Trust," *Proceedings of CHI '01: Extended Abstracts on Human Factors in Computing Systems*, USA: New York.

Cvetkovivh, G. and Lofstedt, R. (1999). *Social Trust and the Management of Risk*. Earthscan Publications.

Damiani, E., Di Vimercati, D., Paraboschi, S., Samarati, P., and Violante, F. (2002). "A Reputation-Based Approach for Choosing Reliable Resources in Peer-to-Peer

Networks," *Proceedings of the 9th ACM conference on Computer and Communications Security (CCS'02)*.

Dash, R., Ramchurn, S., and Jennings, N. (2004). "Trust-Based Mechanism Design," *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*, USA: Washington D.C.

Dawkins, R. (1988). *The Selfish Gene.* Oxford: Oxford University Press.

Dellarocas, C. (2003). "The digitalization of Word-Of-Mouth: Promise and Challenges of Online Reputation Mechanisms," *Management Science* 49(10).

Dellarocas, C. (2000). "Immunizing Online Reputation Reporting Systems Against Unfair Ratings and Discriminatory Behavior," *Proceedings of ACM Conference on Electronic Commerce*.

Demolombe, R. (2001). "To Trust Information Sources: a Proposal for a Modal Logic Framework. Trust and Deception in Virtual Societies," in C. Castelfranchi, and Yao-Hua Tan (Eds.), *Trust and Deception in Virtual Societies*. Kluwer Academic Publisher.

Deutsch, M. (1973). *The Resolution of Conflict*. New Haven: Yale University Press.

Dingledine, R., Freedman, M., and Molnar, D. (2000). *Accountability Measures for Peer-to-Peer Systems, Peer-to-Peer: Harnessing the Power of Disruptive Technologies*. O'Reilly Publishers.

Ding, L., Kolari, P., Finin, T., Joshi, A., Peng, Y., and Yesha, Y. (2005). "On Homeland Security and the Semantic Web: A Provenance and Trust Aware Inference Framework," *Proceedings of the AAAI Spring Symposium on AI Technologies for Homeland Security*.

Ding, L., Kolari, P., Ganjugunte, S., Finin, T., and Joshi, A. (2004). "Modeling and Evaluating Trust Network Inference," *Proceedings of the* 7th *International Workshop on Trust in Agent Societies at AAMAS*.

Ding, L., Zhou, L., and Finin, T. (2003). "Trust Based Knowledge Outsourcing for Semantic Web Agents," *Proceedings of the 2003 IEEE/WIC International Conference on Web Intelligence*.

Doney, M., and Cannon, J. (1997). "An Examination of the Nature of Trust in Buyer–Seller Relationships," *Journalof Marketing* 61, 35–51.

Downey, D., Etzioni, O., and Soderland, S. (2005). "A Probabilistic Model of Redundancy in Information Extraction," *Proceedings of the* 19*th* *International Joint Conference on Artificial Intelligence*.

Dubois, D. and Prade, H. (1980). *Fuzzy Sets and Systems: Theory and Applications*. New York: Academic Press.

eBay (2002). "eBay." Retreived December 11, 2006, from http://www.eBay.com. Checked on January 23, 2007.

EHR Policy (2001). "Electronic health records policy." Retrieved December 11, 2006, from http://www.show.scot.nhs.uk/sehd/publications/DC20011220IMTEHRPol.pdf. Checked on January 23, 2007.

Einwiller, S., Geissler, U., and Will, M. (2000). "Engendering Trust in Internet Businesses Using Elements of Corporate Branding," *Proceedings of the 2000 Americas Conference on Information Systems (AMCIS)*, USA: Long Beach.

Einwiller, S., and Will, M. (2001). "The Role of Reputation to Engender Trust in Electronic Markets," *Proceedings of the Fifth International Conference on Corporate Reputation, Identity, and Competitiveness*, France: Paris.

Ekstrom, M., and Bjornsson, H. (2002). "A Rating System for AEC E-Bidding that Accounts for Rater Credibility," *Proceedings of the CIB W65 Symposium*.

Esfandiari, B., and Chandrasekharan, S. (2001). "On How Agents Make Friends: Mechanisms for Trust Acquisition," *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies,* Canada: Montreal.

Fahrenholtz, D., and Lamesdorf, W. (2002). "Transactional Security for a Distributed Reputation Management System," *Proceedings of the Third International Conference on E-Commerce and Web Technologies (EC-Web)* LNCS 2455, 214–223.

Falcone, R., and Castelfranchi, C. (2001). "Social Trust: a Cognitive Approach," in C. Castelfranchi, and Y. Tan (Eds.), *Trust and Deception in Virtual Societies*. USA, MI: Kluwer Academic Publisher.

Falcone, R., and Castelfranchi, C. (2004). "Trust Dynamics: How Trust is Influenced by Direct Experiences and by Trust Itself," *Proceedings of the Third International*

Joint Conference on Autonomous Agents and Multiagent Systems, USA: Washington D. C.

Falcone, R., Pezzulo, G., and Castelfranci, C. (2003). "A Fuzzy Approach to a Belief-Based Trust Computation," in R. Falcone, S. Barber, L. Korba, and M. Singh (Eds.), *Trust, Reputation, and Security: Theories and Practice.* Springer-Verlag.

Fang, L., Hipel, K., and Kilgour, D. (1993). *Interactive Decision Making: the Graph Model for Conflict Resolution*. New York: Wiley.

Fang, L., Hipel, K., and Kilgour, D. (2003a). "A Decision Support System for Interactive Decision Making, part 1: Model Formulation," *IEEE Transactions on Systems, Man and Cybernetics* 33(C), 42-55.

Fang, L., Hipel, K., and Kilgour, D. (2003b). "A Decision Support System for Interactive Decision Making, part 2: Analysis and Output Interpretation," *IEEE Transactions on Systems, Man and Cybernetics* 33(C), 56-66.

Ferrary, M. (2002). "Trust and Social Capital in the Regulation of Lending Activities," *Journal of Socio-Economics* 235, 1–27.

Fishburn, P. (1988). *Nonlinear Preference and Utility Theory*. Baltimore, MD: Johns Hopkins University Press.

Flake, G. (1998). *The Computational Beauty of Nature.* Cambridge, MA: MIT Press.

Fraser, N., and Hipel, K. (1979). "Solving complex conflicts," *IEEE Transactions on Systems, Man, and Cybernetics* 9, 805-817.

Fraser, N. and Hipel, K. (1984). *Conflict Analysis: Models and Resolutions*. New York: North-Holland.

Freeman, L. (2000). "Centrality on social networks," *Social Networks* 1(1979), 215–239.

Friedman, B., Khan, P., and Howe, D. (2000). "Trust Online," *Communications of the ACM* 43(12), 34–40.

Friedman, E., and Resnick, P. (2001). "The Social Cost of Cheap Pseudonyms," *Journal of Economics and Management Strategy* 10(2), 173–199.

Friedman, M. (1960). *A Program for Monetary Stability*. New York: Fordham University Press.

Fukuyama, F. (1995). *Trust: the Social Virtues and the Creation of Prosperity*. Penguin.

Gambetta, D. (2000). “Can We Trust Trust?,” in D. Gambetta (Eds.), *Trust: Making and Breaking Cooperative Relations*. University of Oxford Press.

Gandon, F., and Sadeh, N. (2004). “Semantic Web Technologies to Reconcile Privacy and Context Awareness,” *Proceedings of the 1st French-Speaking Conference on Mobility and Ubiquity Computing*, USA: New York.

Ganesan, S. (1994). “Determinants of Long-Term Orientation in Buyer–Seller Relationships,” *Journal of Marketing* 58, 1–19.

Ganoulis, J. (1994). *Engineering Risk Analysis of Water Pollution*. VCH and Weinheim.

Gefen, D. (2002). “Reflections on the Dimensions of Trust and Trustworthiness among Online Consumers,” *SIGMIS Database* 33(3), 38–53.

Gefen, D. (2003). “TAM or Just Plain Habit: A Look at Experienced Online Shoppers,” *Journal of End User Computing* 15(3), 1–13.

Gefen, D., Karahanna, E., and Straub, D. (2003). “Trust and TAM in Online Shopping: an Integrated Model,” *MIS Quarterly* **27(**1), 51–90.

Geyskens, I., Steenkamp, J., and Kumar, N. (1997). “Generalizations about Trust in Marketing Channel Relationships Using Meta-Analysis,” *International Journal of Research in Marketing* 15, 223–248.

Gil, Y., and Artz, D. (2006). “Towards Content Trust of Web Resources,” *Proceedings of the* 15*th International World Wide Web Conference*.

Gil, Y., and Ratnakar, V. (2002a). “Trellis: An Interactive Tool for Capturing Information Analysis and Decision Making,” *Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web*, UK: London.

Gil, Y., and Ratnakar, V. (2002b). “Trusting Information Sources One Citizen at a Time,” *Proceedings of the First International Semantic Web Conference on The Semantic Web*, UK: London.

Gill, J., and Butler, R. (1996). “Cycles of Trust and Distrust in Joint-Ventures,” *European Management Journal* 14(1), 81–89.

Glickman, M. (1999). “Parameter Estimation in Large Dynamic Paired Comparison Experiments,” *Applied Statistics* 48, 377–394.

Golbeck, J., and Hendler, J. (2004a). “Accuracy of Metrics for Inferring Trust and Reputation,” *Proceedings of the* 14*th International Conference on Knowledge Engineering and Knowledge Management*.

Golbeck, J., and Hendler, J. (2004b). “Inferring Reputation on the Semantic Web,” *Proceedings of the 13th International World Wide Web Conference*.

Good, D. (2000). “Individuals, Interpersonal Relations, and Trust,” in D. Gambetta (Eds.), *Trust: Making and Breaking Cooperative Relations*. University of Oxford Press.

Govier, T. (1998). *Dilemmas of Trust*. McGill-Queen's University Press.

Grabner-Krauter, S., and Kaluscha, E. (2003). “Empirical Research in On-Line Trust: a Review and Critical Assessment,” *International Journal of Human-Computer Studies* 58, 783-812.

Grandison, T., and Sloman, M. (2000). “A Survey of Trust in Internet Applications,” *IEEE Communications Surveys and Tutorials* 4(4),2–16.

Granovetter, M. (1973). “The Strength of Weak Ties,” *American Journal of Sociology* 78, 1360–1380.

Granovetter, M. (1974). *Getting a Job: A Study of Contacts and Careers*. Harvard: Harvard University Press.

Granovetter, M. (1985). “Economic Action and Social Structure: The Problem of Embeddedness,” *American Journal of Sociology* 91(3), 481–510.

Granovetter, M. (1991). *Society and Economy: The Social Construction of Economic Institutions*. Harvard: Harvard University Press.

Griffiths, N., Chao, K., and Younas, M. (2006). “Fuzzy Trust for Peer-to-Peer Systems,” *Proceedings of the 26th IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW’06),* Portugal: Lisbon, 73-79.

Guha, R., Kumar, R., Raghavan, P., and Tomkins, A. (2004). “Propagation of Trust and Distrust,” *Proceedings of the 13th International Conference on World Wide Web,* USA: New York.

Gupta, M., Judge, P., and Ammar, M. (2003). “A Reputation System for Peerto-Peer Networks,” *Proceedings of the 13th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV).*

Guth, W. (2001). “Do Banks Crowd in Business Ethics? An Indirect Evolutionary Analysis,” *International Review of Economics and Finance* 10(1), 1–17.

Guth, W., and Kliemt, H. (1994). “Competition or Co-operation on the Evolutionary Economics of Trust Exploitation and Moral attitudes,” *Metroeconomica* 45, 155-187.

Guttman, R., Moukas, A., and Maes, P. (1998). “Agent-Mediated Electronic Commerce: a Survey,” *Knowledge Engineering Review* 13(3), 147-159.

Gyongyi, Z., Garcia-Molina, H., and Pedersen, J. (2004). “Combating Web Spam With Trust Rank,” *Technical Report 2004-52*, Stanford.

Haimes, Y. (2004). *Risk Modeling, Assessment, and Management*. Second Edition. New York: Wiley.

Hamouda, L., Kilgour, D., and Hipel, K. (2004). “Strength of Preferences in the Graph Model for Conflict Resolution,” *Group Decision and Negotiation* 13, 449-462.

Hardin, R. (2002). *Trust and Trustworthiness*. Russell Sage Foundation.

Hart, O. (1987). “The Theory of Contracts.” In R. Bewley (Ed.), *Advances in Economic Theory*. Cambridge: Cambridge University Press.

Hayek, F. (1974). *Choice in Currency: A Way to Stop Inflation*. London: The Institute of Economic Affairs.

Hayek, F. (1978). *Desnationalisation of Money: The Argument Refined*. London: The Institute of Economic Affairs, London.

Hipel, K. (2001). “Conflict resolution.” in M. Tolba (Ed.), *Our Fragile World: Challenges and Opportunities for Sustainable Development*. UK: Oxford EOLSS Publishers Co, Ltd.

Hipel, K., Fang, L., Kilgour, D., and Haight, M. (1993). “Environmental Conflict Resolution Using the Graph Model,” *Proceedings of IEEE SMC International Conference*, France: Le Touquet.

Hipel, K., Kilgour, D., Fang, L., and Peng X. (1997). “The Decision Support System GMCR in Environmental Conflict Management,” *Applied Mathematics and Computation* 83(2-3), 117-152.

Hipel, K., and McLeod, I. (1994). *Time series modelling of water resources and environmental systems*. New York : Elsevier.

Hirshleifer, J., and Riley, J. (1979). "The Analytics of Uncertainty and Information: an Expository Survey," *Journal of Economic Literature* 17, 1374–1421.

Hofstadter, D. (1983). "Computer Tournaments of the Prisoner's Dilemma Suggest how Cooperation Evolves," *Scientific American* 248(5), 16-26.

Hood, W., and Wilson, C. (2001). "The Literature of Bibliometrics, Scientometrics, and Informetrics," *Scientometrics* 52(2), 291– 314.

Howard, N. (1971). *Paradoxes of Rationality: Theory of Metagames and Political Behavior*. Cambridge: MIT Press.

Howard, N. (1994). "Drama Theory and its Relationship to Game Theory: I and II," *Group Decision and Negotiation* 3, 187-206.

Howard, N. (1999). *Confrontation Analysis: How to Win Operations Other Than War*. Washington D.C.: CCRP publication series.

Howard, N., Bennett, P., Bryant, J., and Bradley, M. (1992). "Manifesto for a Theory of Drama and Irrational Choice," *Journal of the Operational Research Society* 44, 99-103.

Huberman, B., and Wu, F. (2004). "The Dynamics of Reputations," *Journal of Statistical Mechanics: Theory and Experiment* P04006, 2-17.

Hume, D. (1975). *A Treatise of Human Nature*. Oxford: Clarendon Press.

Huynh, T., Jennings, N., and Shadbolt, N. (2004). "FIRE: An Integrated Trust and Reputation Model for Open Multi-Agent Systems," *Proceedings of the* 16*th European Conference on Artificial Intelligence*.

Inkpen, A., and Currall, S. (1998). "The Nature, Antecedents, and Consequences of Joint Venture Trust," *Journal of International Management* 4(1), 1–20.

Ismail, R., Boyd, C., Jøsang, A., and Russel, S. (2003). "Strong Privacy in Reputation Systems," *Proceedings of the 4th International Workshop on Information Security Applications (WISA)*, Korea: Jeju Island.

iTrust (2005). "iTrust." Retrieved December 20, 2006, from http://www.itrust.uoc.gr/. Checked on January 23, 2007.

Jensen, M., and Meckling,W. (1976). "Theory of the Firm: Managerial Behavior, Agency Cost, and Ownership Structure," *Journal of Financial Economics* 3(4), 305-360.

Jones, A. (2002). "On the Concept of Trust," *Decision Support Systems* 33, 225–232.

Jonker, C., Schalken, J., Theeuwes, J., and Treur, J. (2004). "Human Experiments in Trust Dynamics," *Lecture Notes in Computer Science* 2995, 206–220.

Jøsang, A. (1999). "Trust-Based Decision Making for Electronic Transactions," *Proceedings of the 4th Nordic Workshop on Secure Computer Systems (NORDSEC'99)*, Sweden: Stockholm University.

Jøsang, A. (2001). "A Logic for Uncertain Probabilities," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 9(3), 279– 311.

Jøsang, A., and Ismail, R. (2002). "The Beta Reputation System," *Proceedings of the 15th Bled Electronic Commerce Conference*, Slovenia: Bled.

Jøsang, A., and Presti, S. (2004). "Analysing the Relationship between Risk and Trust," *Proceedings of the Second International Conference on Trust Management*, UK: Oxford.

Jøsang, A., and Pope, S. (2005). "Semantic Constraints for Trust Transitivity," *Proceedings of the Asia-Pacific Conference of Conceptual Modeling (APCCM)*, Australia: Newcastle.

Jøsang, A., Hird, S., and Faccer, E. (2003). "Simulating the Effect of Reputation Systems on E-Markets," *Proceedings of the First International Conference on Trust Management*, Crete.

Jurca, R., and Faltings, B. (2003). "An Incentive Compatible Reputation Mechanism," *Proceedings of the 6th Int. Workshop on Deception Fraud and Trust in Agent Societies (at AAMAS'03)*.

Kamvar, S., Schlosser, M., and Garcia-Molina, H. (2003). "The Eigen-Trust Algorithm for Reputation Management in P2P Networks," *Proceedings of the Twelfth International World Wide Web Conference*, Budapest.

Kandasamy, W. and Smarandache, F. (2003). *Fuzzy Cognitive Maps and Neutrosophic Cognitive Maps*. Publishing Online Co.

Kandel, A. and Zhang, Y. (1998). "Fuzzy Moves," *Fuzzy Sets and Systems*, 99, 159-177.

Karlins, M., and Abelson, H. (1970). *Persuasion, How Opinion and Attitudes are Changed*. Crosby Lockwood & Son.

Karwowski, W. and Mital, A. (1986). “Applications of Approximate Reasoning in Risk Analysis,” in G. Klir, U. Clair, and B. Yuan (Eds.), *Fuzzy Set Theory*: *Foundations and Applications*. Amsterdam: Elsevier.

Kaufman, A., and Gupta, M. (1985). *Introduction to Fuzzy Arithmetic*, *Theory and Applications*. New York: Van Nostrand.

Khare, R., and Rifkin, A. (1997). “Weaving a web of trust,” *Journal of the World Wide Web* 2(3), 77–112.

Kilgour, D. (1984). “Equilibria for Far-Sighted Players,” *Theory and Decision* 16, 135-157.

Kilgour, D. (1985). “Anticipation and Stability in Two-Person Nonooperative games.” in M. Ward and U. Luterbacher (Eds.), *Dynamic Models of Inter- national Conflict*, Boulder, Colorado: Lynne Rienner Press.

Kilgour, D., Fang, L., Hipel, K. (1995). “GMCR in Negotiations,” *Negotiation Journal* 11(2), 151-156.

Kilgour, D., K. Hipel, and Fang, L. (2001). “Coalition Analysis in Group Decision Support,” *Group Decision and Negotiation* 10, 159-175.

Kim, S., and Hovy, E. (2004). “Determining the Sentiment of Opinions,” *Proceedings of the* 20*th International Conference on Computational Linguistics (COLING)*.

Klein, B. (1974). “The Competitive Supply of Money,” *Journal of Money Credit and Banking* November, 423-53.

Kleinberg, J. (1999). “Authoritative Sources in a Hyperlinked Environment,” *Journal of the ACM* 46(5), 604–632.

Kohl, J., and Neuman, B. (1994). “The Kerberos Network Authentication Service,” *IEEE Communications* 32(9), 33-38

Kollock, P. (1993). “An Eye for an Eye Leaves Everyone Blind: Cooperation and Accounting Systems,” *American Sociological Review* 58, 768-786.

Kosko, B. (1986). “Fuzzy Cognitive Maps,” *International Journal of Man-Machine Studies* 24, 65-75.

Kraines, D., and Kraines, V. (1988). “Pavlov and the Prisoner's Dilemma,” *Theory and Decision* 26, 47-79.

Kreps, D., and Wilson, R. (1982). “Reputation and Imperfect Information,” *Journal of Economic Theory* 27(2), 253– 279.

Kurbel, I., and Loutchko, K. (2001). “A Framework for Multi-Agent Electronic Marketplaces: Analysis and Classification of Existing Systems,” *Proceedings of the International ICSC Congress on Information Science Innovations (ISI’01)*, UAE: American University in Dubai.

Lahno, B. (2001). “On The Emotional Character of Trust,” *Ethical Theory and Moral Practice* 4, 171-189.

Lashkari, Y., Metral, M., and Maes, P. (1994). “Collaborative Interface Agents,” *Proceedings of the Twelfth National Conference on Artificial Intelligence*.

Leithead, T., Nejdl, W., Olmedilla, D., Seamons, K., Winslett, M., Yu, T., and Zhang, C. (2004). “How to Exploit Ontologies for Trust Negotiation,” *Proceedings of ISWC Workshop on Trust, Security, and Reputation on the Semantic Web*, Japan: Hiroshima.

Levien, R. (2004). *Attack Resistant Trust Metrics*. PhD thesis. USA: University of California at Berkeley.

Lewicki, R., and Bunker, B. (1995). “Trust in Relationships: A Model of Trust Development and Decline.” in B. Bunker, and J. Rubin (Eds.), *Conflict, Cooperation and Justice*. San Francisco: Jossey- Bass.

Li, K., Hipel, K., Kilgour, D., and Fang, L. (2004). “Preference Uncertainty in the Graph Model for Conflict Resolution,” *IEEE Transactions on Systems, Man, and Cybernetics Part A* 34(4), 507-520.

Li, K., Karray, F., Hipel, K., and Kilgour, D. (2001). “Fuzzy Approaches to the Game of Chicken,” *IEEE Transactions on Fuzzy Systems* 9, 608-623.

Li, Y., Huang, G., Nie, S., and Huang, Y. (2006). “IFTSIP: Interval Fuzzy Two-Stage Stochastic Mixed-Integer Programming: a Case Study for Environmental Management and Planning,” *Civil Engineering and Environmental Systems* 23(2), 73-99.

Liau, C., Zhou, X., Bressan, S., and Tan, K. (2003). “Efficient Distributed Reputation Scheme for Peer-to-Peer Systems,” *Proceedings of the 2nd International Human.Society@Internet Conference (HSI)* LNCS 2713, pp. 54–63.

Lin, C. and Chen, C. (2004). "a Fuzzy-Logic-Based Approach for New Product Go/NoGo Decision at the Front End," *IEEE Transactions on Systems, Man and Cybernetics* 34(1), 132-142.

Lin, C. and Chen, Y. (2004). "Bid/No-Bid Decision-Making – a Fuzzy Linguistic Approach," *International Journal of Project Management* 22, 585-593.

Littlestone, N., and Warmuth, M. (1994). "The Weighted Majority Algorithm," *Information and Computation* 108(2), 212– 261.

Lubell, M. (2001, December). "Trust and Cooperation in Watershed Management." Retrieved Spetember 23, 2004, from http://www.des.ucdavis.edu/faculty/lubell/Research/GrantRevisionFinal.pdf. Checked on January 23, 2007.

Luhmann, N. (1979). *Trust and Power*. New York: Wiley.

Luhmann, N. (2000). "Familiarity, Confidence, Trust: Problems and Alternatives," in D. Gambetta (Eds.), *Trust: Making and Breaking Cooperative Relations*. University of Oxford Press.

Luo, Y. (2002). "Building Trust in Cross-Cultural Collaborations: Toward a Contingency Perspective," *Journal of Management* 28(5), 669–694.

Maes, P., Guttman, R., and Moukas, A. (1999). "Agents that Buy and Sell: Transforming Commerce as We Know It," *Communications of the ACM* 42(3), 81–91.

Makino, J. (1988). "Productivity of Research Groups - Relation between Citation Analysis and Reputation within Research Communities," *Scientometrics* 43(1), 87-93.

Manchala, D. (1998). "Trust Metrics, Models and Protocols for Electronic Commerce Transactions," *Proceedings of the 18th International Conference on Distributed Computing Systems*, 1998.

Mares, M. (2001). *Fuzzy Cooperative Games, Cooperation with Vague Expectations*. Germany: Physica-Verlag.

Marimon, R., Nicolini, J., and Teles, P. (2000). "Competition and Reputation," *Proceedings of the World Conference Econometric Society,* USA: Seattle.

Marsden, P., and Lin, N. (1982). *Social Structure and Network Analysis*, Beverly Hills: Sage Publications, 1982.

Marsh, S. (1994). *Formalising Trust as a Computational Concept*. PhD thesis, University of Stirling.

Massa, P., and Avesani, P. (2005). "Controversial Users Demand Local Trust Metrics: an Experimental Study on Epinions.com Community," *Proceedsings of the* 25*th American Association for Artificial Intelligence Conference*.

Massa, P., and Hayes, C. (2005). "Page-re-Rank: Using Trusted Links to Re-Rank Authority," *Proceedings of the The 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI'05)*, USA: Washington D.C.

Maximilien, E., and Singh, M. (2005). "Agent Based Trust Model Involving Multiple Qualities," *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'05)*, Netherlands: Utrecht, 519-526.

Mayer, R., Davis, J., and Schoorman, F. (1995). "An Integrative Model of Organizational Trust," *Academy of Management Review* 20(3), 709–734.

McGuinness, D. (2004). "Question Answering on the Semantic Web," *In IEEE Intelligent Systems*, 19(1), 82-85.

McKnight, D., and Chervany, N. (1996). "The Meanings of Trust." Retrieved October 23, 2006, from http://misrc.umn.edu/wpaper/. Checked on January 23, 2007.

McKnight, D., and Chervany, N. (2002). "Notions of Reputation in Multi-Agent Systems: A Review," *Proceedings of the 34th Hawaii International Conference on System Sciences*.

McKnight, D., Cummings, L., and Chervany, N. (1998). "Initial Trust Formation in New Organizational Relationships," *Academy of Management Review* 23(3), 473–490.

Mesterton-Gibbons, M., and Dugatkin, L. (1992). "Cooperation among Unrelated Individuals: Evolutionary Factors," *Quarterly Review of Biology* 67, 267-281.

Miller, G. (1965). "The Magic Number Seven, Plus or Minus Seven," Psychology Review 63, 81–97.

Miller, N., Resnick, P., and Zeckhauser, R. (2002). “Eliciting Honest Feedback in Electronic Markets.” Retrieved October 12, 2006 from http://www.si.umich.edu/presnick/papers/elicit/. Checked January 23, 2007.

Molander, P. (1985). “The Optimal Level of Generosity in a Selfish, Uncertain Environment,” *Journal of Conflict Resolution* 29, 611-618.

Montaner, M., Lopez, B., and de la Rosa, J. (2002). “Developing Turst in Recommender Agents,” *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02),* Italy: Bologna.

Morgan, R., and Hunt, S. (1994). “The Commitment–Trust Theory of Relationship Marketing,” *Journal of Marketing* 58, 20–38.

Mui, L., Mohtashemi, M., Ang, C., Szolovits, P., and Halberstadt, A. (2001). “Ratings in Distributed Systems: a Bayesian Approach,” *Proceedings of the Workshop on Information Technologies and Systems (WITS)*.

Mui, L., Halberstadt, A., and Mohtashemi, M. (2002). “Notions of Reputation in Multi-Agent Systems: A Review,” *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02),* Italy: Bologna.

Mui, L., Mohtashemi, M., and Ang, C. (2001). “A Probabilistic Rating Framework for Pervasive Computing Environments,” *Proceedings of the MIT Student Oxygen Workshop (SOW’2001)*.

Mui, L., Mohtashemi, M., and Halberstadt, A. (2002). “A Computational Model of Trust and Reputation,” *Proceedings of the 35th Hawaii International Conference on System Science (HICSS)*.

Mukherji, A., Rajan, V., and Slagle, J. (1996). “Robustness of Cooperation,” *Nature* 379, 125-126.

Nash, J. (1950). “Equilibrium Points in N-Person Games,” *Proceedings of National Academy of Science*.

Nash, J. (1951). “Non-Cooperative Games,” *Annals of Mathematics* 54, 268-295.

Nejdl, W., Olmedilla, D., and Winslett, M. (2004). “Peertrust: Automated Trust Negotiation for Peers on the Semantic Web,” *Proceedings of Workshop on Secure*

Data Management in a Connected World in conjunction with the 30*th International Conference on Very Large Data Bases*.

Neumann, J. V. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton: University Press.

Nguyen, H. (1997). *A First Course in Fuzzy Logic*. CRC Press.

Nielsen, M., and Krukow, K. (2003). “Towards a Formal Notion of Trust,” *Proceedings of the 5th ACM SIGPLAN international conference on Principles and practice of declaritive programming* USA: New York.

Nowak, M., and May, R. (1992). “Evolutionary Games and Spatial Chaos,” *Nature* 359, 826-829.

Nowak, M., and Sigmund, K. (1993b). “A Strategy of Win-stay, Lose-shift that Outperforms Tit-for-tat in the Prisoner's Dilemma Game,” *Nature* 364, 56-58.

Nowak, M., and Sigmund, K. (1992). “Tit for Tat in Heterogeneous Populations,” *Nature* 355, 250-252.

O’Donovan, J., and Smyth, B. (2005). “Trust in Recommender Systems,” *Proceedings of the 10th international conference on Intelligent user interfaces* USA: New York.

O’Hara, K., Alani, H., Kalfoglou, Y., and Shadbolt, N. (2004). “Trust Strategies for the Semantic Web,” *Proceedings of Workshop on Trust, Security, and Reputation on the Semantic Web,* 3rd *International Semantic Web Conference*.

OnSale (2002). “OnSale.” Retrieved December 12, 2006 from http://www.onsale.com. Checked Jauary 23, 2007.

Osborne, M. (2003). *An Introduction to game Theory*. New York: Oxford University Press.

Page, L., Brin, S., Motwani, R., and Winograd, T. (1998). “The PageRank Citation Ranking: Bringing Order to the Web.” Retrieved January 2, 2007 from http://dbpubs.stanford.edu:8090/pub/1999-66. Checked January 23, 2007.

Pavlou, P. (2003). “Consumer Acceptance of Electronic Commerce — Integrating Trust and Risk with the Technology Acceptance Model,” *International Journal of Electronic Commerce* 7(3), 69–103.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.

Peng, X. (1999). *A Decision Support System for Conflict Resolution*. PhD thesis, University of Waterloo, Waterloo, Ontario, Canada.

Pirzada, A., and McDonald, C. (2004). "Establishing Trust in Pure Adhoc Networks," *Proceedings of the 27th conference on Australasian computer science*, Australia: Darlinghurst.

Plato (1955). *The Republic (370BC)*. Viking Press.

Potter, N. (2002). *How Can I Be Trusted: a Virtue of Trustworthiness*. Rowman & Littlefield.

Ramchurn, S., Sierra, C., Godo, L., and Jennings, N. (2003). "A Computational Trust Model for Multi-Agent Interactions Based on Confidence and Reputation," *Proceedings of the* 6*th International Workshop of Deception, Fraud and Trust in Agent Societies*.

Rapaport, A., and Chammah, A. (1965). *Prisoner's Dilemma*. Ann Arbor , MI: University of Michigan Press.

Rasmusson, L., and Janssen, S. (1996). "Simulated Social Control for Secure Internet Commerce," *Proceedings of the 1996 New Security Paradigms Workshop*.

Reiter, M., and Stubblebine, S. (1997). "Toward Acceptable Metrics of Authentication," *Proceedings of the 1997 IEEE Symposium on Research in Security and Privacy*, USA: Oakland.

Resnick, P., and Zeckhauser, R. (2002). "Trust among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System," in M. Baye (Ed.), *The Economics of the Internet and ECommerce, Advances in Applied Microeconomics*. Elsevier Science.

Resnick, P., Zeckhauser, R., Friedman, R., and Kuwabara, K. (2000). "Reputation Systems," *Communications of the ACM* 43(12), 45–48.

Resnick, P., Zeckhauser, R., Swanson, J., and Lockwood, K. (2002). "The Value of Reputation on eBay: a Controlled Experiment," Retrieved December 22, 2006 from http://www.si.umich.edu/presnick/papers/postcards/. Checked on January 23, 2007.

Richardson, M., Argawal, R., and Domingos, P. (2003). "Trust Management for the Semantic Web," *Proceedings of the Second International Semantic Web Conference*.

Riggs, T., and Wilensky, R. (2001). "An Algorithm for Automated Rating of Reviewers," *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL)*.

Riloff, E., Wiebe, J., and Phillips, W. (2005). "Exploiting Subjectivity Classification to Improve Information Extraction," *Proceedings of the* 20th *National Conference on Artificial Intelligence*.

Rockoff, H. (1975). *The Free Banking Era: A Reconsideration*. New York: Arno Press.

Rousseau, D., Sitkin, S., Burt, R., and Camere, C. (1998). "Not so Different after All: a Cross-Discipline View of Trust," *Academy of Management Review* 23(3), 393-404.

Rotter, J. (1967). "A New Scale for the Measurement of Interpersonal Trust," *Journal of Personality* 35, 651-665.

Rotter, J. (1971). "Generalized Expectancies for Interpersonal Trust," *American Psychologist* 26(5), 443–452.

Rutter, J. (2001). "From the Sociology of Trust Towards a Sociology of 'E-Trust'," *International Journal of New Product Development and Innovation Management* 2(4), 371-385.

Sabater, J., and Sierra, C. (2001). "REGRET: a Reputation Model for Gregarious Societies," *Proceedings of the 4th International Workshop on Deception, Fraud and Trust in Agent Societies, in the 5th International Conference on Autonomous Agents (AGENTS'01)*, Canada: Montreal.

Sabater, J., and Sierra, C. (2002). "Reputation and Social Network Analysis in Multi-Agent Systems," *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

Sabater, J., and Sierra, C. (2002). "Social ReGreT, a Reputation Model Based on Social Relations," *ACM SIGecom Exchanges* 3(1), 44–56.

Sakakibara, H., Okada, N., and Nakase, D. (2002). "The Application of Robustness Analysis to the Conflict with Incomplete Information," *IEEE Transactions on Systems, Man, and Cybernetics Part C: Applications and Reviews* 32(1), 14-23.

SAML, (2005). "SAML." Retrieved December 21, 2006, from http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=security. Checked on January 23, 2007.

Schiff, A., and Kennes, J. (2003). "The Value of Reputation Systems," *Proceedings of the First Summer Workshop in Industrial Organization (SWIO)*, NZ: Auckland.

Schillo, M., Funk, P., and Rovatsos, M. (2000). "Using Trust for Detecting Deceitful Agents in Artificial Societites," *Applied Artificial Intelligence* (Special Issue on Trust, Deception and Fraud in Agent Societies), 825-848.

Schmucker, K. (1985). *Fuzzy Sets, Natural Language Computations, and Risk Analysis*. Maryland: Computer Science Press.

Schneider, J., Kortuem, G., Jager, J., Fickas, S., and Segall, Z. (2000). "Disseminating Trust Information in Wearable Communities," *Proceedings of the 2nd International Symposium on Handheld and Ubiquitous Computing (HUC2K)*.

Scott, J. (2000). *Social Network Analysis*. SAGE Publications.

Sen, S., and Sajja, N. (2002). "Robustness of Reputation-Based Trust: Boolean Case," *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

Sen, S., Sekaran, M., and Hale, J. (1994). "Learning To Coordinate Without Sharing Information," *Proceedings of the 12th National Conference on Artificial Intelligence*.

Seigneur, J., and Jensen, C. (2004). "Trust Enhanced Ubiquitous Payment Without too Much Privacy loss," *Proceedings of the 2004 ACM symposium on Applied computing*, USA: New York.

Shankar, V., Sultan, F., Urban, G., and Bart, I. (2002). "The Role of Trust in Online Customer Support," *Working Paper*, Sloan School of Management, MIT, Cambridge, MA 02142.

Shapiro, C. (1982). "Consumer Information, Product Quality, and Seller Reputation," *The Bell Journal of Economics* 13(1), 20– 35.

Shapiro, C. (1983). "Premiums for High Quality Products as Returns to Reputations," *Quarterly Journal of Economics* 48, 659-680.

Shapiro, C., and Stiglitz, J. (1984). "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review* 74, 433–444.

Sigmund, K. (1993). *Games of Life.* Oxford: Oxford University Press, 1993.

Sillence, E., Briggs, P., Fishwick, L., and Harris, P. (2004). "Trust and Mistrust of Online Health Sites," *Proceedings of the SIGCHI conference on Human factors in computing systems*, USA: New York.

Simmhan, Y., Plale, B., and Gannon, D. (2005). "A Survey of Data Provenance in E-Science," *Special Interest Group on Management of Data Record* 34(3), 31–36.

Skyrms, Brian. (2003) *The stag hunt and Evolution of Social Structure*. Cambridge: Cambridge University Press.

Smith, J. (1982). *Evolution and the Theory of Games.* Cambridge: Cambridge University Press.

Song, Q. and Kandel, A. (1999). "A Fuzzy Approach to Strategic Games," *IEEE Transactions on Fuzzy Systems* 7(6), 634-642.

Staab, S. et al. (2004). "The Pudding of Trust," *IEEE Intelligent Systems* 19(5), 74–88.

Stephens, R. (2004). "A Framework for the Identification of Electronic Commerce Design Elements that Enable Trust within the Small Hotel Industry," *Proceedings of the 42nd annual Southeast regional conference*, USA: New York.

Steuer, R. (1986). *Multiple Criteria Optimization: Theory, Computation and Application*. USA: Wiley.

Stewart, K. (1999). "Transference as a Means of Building Trust in World Wide Web Sites," *Proceeding of the 20th international conference on Information Systems*, USA: Atlanta.

Stewart, K., and Zhang, Y. (2003). "Effects of Hypertext Links on Trust Transfer," *Proceedings of the 5th international conference on Electronic commerce*, USA: New York.

Stokey, N. (1991). "Credible Public Policy," *Journal of Economic Dynamics and Control* 15, 627-656.

Stoyanov, V., Cardie, C., and Wiebe, J. (2005). "Multi-Perspective Question Answering Using the Opqa Corpus," *Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language*.

Swan, J., Bowers, M., and Richardson, L. (1999). "Customer Trust in the Salesperson: an Integrative Review and Meta-Analysis of the Empirical Literature," *Journal of Business Research* 44, 93–107.

Sztompka, P. (1999). *Trust: A Sociological Theory*. Cambridge University Press.

Tadelis, S. (2003). "Firm Reputation with Hidden Information," *Economic Theory* 21(2), 635– 651.

Tanino, T. (1988). "Fuzzy Preference Relations in Group Decision Making," in J. Kacprzyk and M. Roubens (Eds.), *Non- Conventional Preference Relations in Decision Making*. Berlin: Springer-Verlag.

Tanino, T. (1990). "On Group Decision Making Under Fuzzy Preferences," in J. Kacprzyk and M. Fedrizzi (Eds.), *Multiperson Decision Making Using Fuzzy Sets and Possibility Theory*. Netherlands: Kluwer Academic Publishers.

Taub, B. (1985). "Private Fiat Money with Many Suppliers," *Journal of Monetary Economics* 16, 195–208.

Taub, B. (1986). "Asymptotic Properties of Pipeline Control of the Money Supply," *International Economic Review* 27(3), 647-665.

Tomkins, C. (2001). "Interdependencies, Trust and Information in Relationships, Alliances and Networks," *Accounting, Organizations and Society* 26, 161–191.

Tonti, G., Bradshaw, J., Jeffers, R., Montanari, R., Suri, N., and Uszok, A. (2003). "Semantic Web Languages for Policy Representation and Reasoning: A Comparison of KAoS, Rei, and Ponder," *Proceedings of the 2003 International Semantic Web Conference*.

Tuomela, M. and Hofman, S. (2003). "Simulating Rational Social Normative Trust, Predictive Trust, and Predictive Reliance between Agents," *Ethics and Information Technology* 5, 163-176.

Viega, J., Kohno, T., and Potter, B. (2001). "Trust (and Mistrust) in Secure Applications," *Communications of ACM* 44(2), 31–36.

Wang, M., Hipel, K. W., and Fraser, N. M. (1988). "Modeling Misperception in Games," *Behavioral Science* 33(3), 207-223.

Warren, R., Johnson, J., and Huang, G. (2004). "Application of Rough Sets to Environmental Engineering Models," *Transactions on Rough Sets* 3100, 356-374.

Wicks, A., Berman, S., and Jones, T. (1999). "The Structure of Optimal Trust: Moral and Strategic Implications," *Academy of Management Review* 24(1), 99–116.

Williamson, O. (1993). "Calculativeness, Trust and Economic Organization," *Journal of Law and Economics* 36, 453– 486.

Winsborough, W., Seamons, K., and Jones, V. (2000). "Automated Trust Negotiation," *Technical Report*, North Carolina State University at Raleigh, Raleigh, NC, USA.

Winslett, M., Yu, T., Seamons, K., Hess, A., Jacobson, J., Jarvis, R., Smith, B., and Yu, L. (2002). "Negotiating Trust on the Web," *IEEE Internet Computing* 6(6), 30–37.

Withby, A., Jøsang, A., and Indulska, J. (2004). "Filtering out Unfair Ratings in Bayesian Reputation Systems," *Proceedings of the 7th International Workshop on Trust in Agent Societies (at AAMAS'04)*.

Wong, S., Miles, S., Fang, W., Groth, P., and Moreau, L. (2005). "Provenance-Based Validation of E-Science Experiments," *Proceedings of the 4th International Semantic Web Conference* 3729, 801–815.

WS-Trust (2005). "WS-Trust." Retrieved December 21, 2006, from http://www-128.ibm.com/developerworks/library/specification/ws-trust/. Checked on January 2007.

XACML (2005). "XACML." Retrieved December 21, 2006, from http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml. Checked on January 23, 2007.

Xiao, S., and Benbasat, I. (2003). "The Formation of Trust and Distrust in Recommendation Agents in Repeated Interactions: a Process-Tracing Analysis," *Proceedings of the 5th international conference on Electronic commerce*, USA: New York.

Yin, Y., Huang, G., and Hipel, K. (1999). "Application of Fuzzy Relation Analysis to Multi-Criteria Water Resources Management," *Journal of Water Resources Planning and Management* 125, 41-47.

Yu, B., and Singh, M. (2000). "A Social Mechanism of Reputation Management in Electronic Communities," *Proceedings of the 4th International Workshop on Cooperative Information Agents*.

Yu, B., and Singh, M. (2001). "Towards a Probabilistic Model of Distributed Reputation Management," *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies,* Canada: Montreal.

Yu, B., and Singh, M. (2002). "An Evidential Model of Distributed Reputation Management," *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS).*

Yu, B., and Singh, M. (2002a). "Distributed Reputation Management for Electronic Commerce," *Computational Intelligence* 18(4), 535–549.

Yu, B., and Singh, M. (2002b). "An Evidential Model of Distributed Reputation Management," *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02),* Italy: Bologna.

Yu, B., and Singh, M. (2003). "Detecting Deception in Reputation Management," *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS).*

Yu, T., and Winslett, M. (2003). "Policy Migration for Sensitive Credentials in Trust Negotiation," *Proceedings of the 2003 ACM workshop on Privacy in the electronic society*, USA: New York.

Yu, T., Winslett, M., and Seamons, K. (2001). "Interoperable Strategies in Automated Trust Negotiation," *Proceedings of the 8th ACM conference on Computer and Communications Security*, USA: New York.

Zacharia, G., Moukas, A., and Maes, P. (1999). "Collaborative Reputation Mechanisms in Electronic Marketplaces," *Proceedings of the 32nd Hawaii International Conference on System Science*.

Zacharia, G. and Maes, P. (2000). "Trust Management through Reputation Mechanisms," *Applied Artificial Intelligence* 14, 881-907.

Zadeh, L. (1973). "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," *IEEE Transactions Systems, Man and Cybernetics*, 3, 28-44.

Zagare, F. (1984). "Limited-Move Equilibria in 2 x 2 Games," *Theory and Decision* 22, 1-19.

Zeckhauser, R. and Viscusi, W. (1990). "Risk within Reason," *Science* 248, 559-564.

Zhao, J., Wroe, C., Goble, C., Stevens, R., Quan, D., and Greenwood, M. (2004). "Using Semantic Web Technologies for Representing E-Science Provenance," *Proceedings of the* 3*rd International Semantic Web Conference*.

Zheng, J., Veinott, E., Bos, N, Olson, J., and Olson, G. (2002). "Trust Without Touch: Jumpstarting Long-Distance Trust with Initial Social Activities," *Proceedings of the SIGCHI conference on Human factors in computing systems*, USA: New York.

Ziegler, C., and Lausen, G. (2004). "Spreading Activation Models for Trust Propagation," *Proceedings of the IEEE International Conference on e-Technology, e-Commerce, and e-Service (EEE '04)*, Taiwan: Taipei.

Zimmermann, H. (1987). *Fuzzy Sets, Decision Making, and Expert Systems*. Boston: Kluwer Academic Publishers.

Zimmermann, P. (1995). *The Official PGP User's Guide*, MIT Press.

Zucker, L. (1986). "Production of Trust: Institutional Sources of Economic Structure, 1840–1920," *Research in Organizational Behavior* 6, 53–111.