Empirical Studies in Computer-Mediated Interest-Based Negotiations

Sohan D'souza

Master of Science in Information Technology
Faculty of Informatics
The British University in Dubai

April 8, 2009

Abstract

Negotiations in which participants exchange offers based on their chosen positions can be extended to include dialogue about their interests. Revelation of negotiators' interests allows them to make more acceptable offers and perhaps propose possible alternative approaches toward each other's interests, both of which may result in mutually and individually beneficial outcomes. However, it can also expose their strategies, and possibly their dependencies on other negotiators toward the achievement of their goals. Revealing this information can leave them vulnerable to extortion or retribution, but it can also be used to gain sympathy or build a relationship of trust and reciprocity.

This dissertation studies human behaviour and performance upon introducing options for goal inquiry and revelation into mediated-protocol negotiation scenarios. Empirical studies were conducted by having human players negotiate over an alternating offer protocol and an interest-based bargaining protocol, on a platform specially adapted for this purpose. The analysis of data from these experiments revealed interesting patterns in the human use of goal revelation, and its effects on individual and social outcomes and likelihood of agreement. The design of the experiments and the development of the experimentation platform lay the groundwork for the further study of goal revelation in mediated negotiations with humans.

Acknowledgement

Dedication

To my father and my mother, for their encouragement and support

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

<div align="right">

(*Sohan D'souza*)

</div>

Contents

List of Figures

List of Tables

Chapter 1

Overview

This chapter introduces the academic context of the dissertation, and summarizes what was achieved, how it was achieved, and how this achievement advances the state of the art in the relevant discipline. It defines the aims and scope of the dissertation, and outlines the organization therof.

Section 1.1 introduces the context of the dissertation, followed by Section 1.2, where the dissertation' research questions are listed, and Section 1.3, where the methodology of study used to answer those questions is explained. Section 1.4 defines the scope of the dissertation, and Sections 1.5 and 1.6 outline the findings of the study and the significance of those findings, respectively. Finally, Section 1.7 describes the structure of the remainder of the document.

1.1 Introduction

A process fundamental to economic activity, negotiation is rich with potential areas of academic exploration and commercial significance for economists, informaticians and psychologists alike. Those involved in the study of multi-agent systems study negotiation so they can improve the effectiveness of their systems in distributing resources and creating value. Those involved in decision support systems can use the knowledge derived from negotiation studies to build negotiation support applications that work in the real world in the face of the complexities of economic dynamics and human psychology. Psychologists study negotiation in order to understand in a more formal way how humans think when solving individual or social problems.

A complex negotiation problem, especially one featuring multiple attributes, may be difficult to resolve by humans. This difficulty may be exacerbated when a negotiation problem arises regularly and frequently, and having a human resolve it each time on time becomes prohibitively tedious and inefficient. On the other hand, due to the large space of possible deals to explore and evaluate, and in situations of incomplete information, computers may also find it difficult to handle negotiations in a reasonable amount of time.

When the parties being negotiated with are human, and the negotiations are to take place in a large-scale e-commerce context, dealing with the large number of negotiations against human participants that result would require some form of assistance or automation. E-commerce contexts, and B2B contexts in particular, are dynamic and characterized by multiple contract attributes, fluid pricing, and fluid back-and-forth negotiation between buyers and sellers. Negotiation support tools could create opportunities to not only handle price wars but determine sellers and buyers preferences across multiple issues and terms (Goh et al., 2000). Should such tools be empowered with heuristics and training derived from the behaviour and performance of human negotiators, they would be invaluable in those cases when they have to face human negotiators.

As with any conflict, information is a strategic asset. It can be used to gain the upper hand

or to gain sympathy, to project trust or to project power, to exploit or to aid. This work studies the trade-offs associated with two different negotiation protocols in negotiation settings where self-interested parties lack information about each others interests.

In one protocol, participants alternate between making take-it-or-leave-it deals to each other under time constraints. A penalty is placed on rejected offers, in order to have participants carefully consider both their proposals as well as their responses. As they cannot reveal their interests to other participants, the offers of participants serve as a noisy signal to their true objectives. It could be difficult to locate efficient trades for both parties in such conditions, either because participants do not know what resources their opponents need, if any, or because there are simply too many combinations of possible agreements to try out under time constraints. Therefore, they may repeatedly make offers that focus on the most beneficial route to their goals (usually the shortest), while ignoring other routes that may be less intuitively beneficial, but for which their opponents may be able and willing to transfer resources.

In these conditions, revealing the objectives of one or more of the participants may facilitate agreement, because the additional information narrows the "search space" of possible offers, and may reveal new avenues of negotiation that were not known before. However, it is not a given that the revelation of objective information by either party will necessarily improve the result of the negotiation. Having an opponent that knows one's goals can be useful but, depending on one's position, potentially costly. This is because it exposes the negotiating position and the negotiation strategies that revealing parties are using. For example, a dependent player may be extorted to give away more chips than it requests in exchange for the ones it minimally requires. A participant that revealed its goals could have exposed itself to be holding back and misrepresenting its position in its past offers, and this may result in retribution in future offers or even a collapse of the negotiation. On the other hand, exposing one's desperate situation may induce charity in the party to whom the knowledge has been made available.

The second protocol therefore allows negotiators to ask for information about their opponents' interests, and allows a negotiator from whom this information has been solicited to choose to reveal it. In order to retain the advantages in synchronization and simplicity that turn-based play offers, these new dialogues are integrated into the flow of the alternating offer protocol described earlier in this section. By studying the difference in performance and behaviour of players with different task dependencies across a number of negotiation scenarios, the research questions posed next will be answered.

Figures 1.1, 1.2 and 1.3 depict an example that illustrates how goal revelation can affect a negotiation. In the scenario imagined here, Alice and Bob are two rival cooks at a sandwich bar. Suppose they start out with the following sandwich fillings allotted to them, and with culinary capabilities described as follows.

Alice has to herself some peanut butter, a pickle spear, some bacon strips, a tomato and cheddar squares, and she has the skill to prepare either of two classic sandwiches: the BLT (bacon-lettuce-tomato) or the PB&J (peanut butter and jelly spread). Bob has a patty of hamburger, a leaf of lettuce, rings of onion, some cream cheese and some jelly spread, and can prepare basic cheeseburger sandwiches of hamburger with pickles and cheddar or sandwiches of cream cheese with jelly spread (CC&J). Neither Alice nor Bob knows what the other cook can prepare, or knows of valid sandwich combinations other than what they can prepare.

Suppose that the manager has asked for a sandwich from each cook for the day's sampling, and has requested them to be conservative with the ingredients, as always. On the outset, Alice does not have the ingredients to make a sandwich that she knows to prepare. Since the least expensive sandwich she can make is the PB&J, she attempts to trade each of her ingredients (other than peanut butter, of course) for Bob's jelly spread. However, since he has the ingredients to make a sandwich - the CC&J - he will not trade his jelly spread (Figure 1.1). For that matter, Bob will not even counter-propose for the pickle and cheddar, as the CC&J requires fewer ingredients than the cheeseburger.

If, however, Bob asks what kind of sandwiches Alice can make, and Alice reveals this information (Figure 1.2), Bob can immediately propose a trade of his lettuce for Alice's cheddar
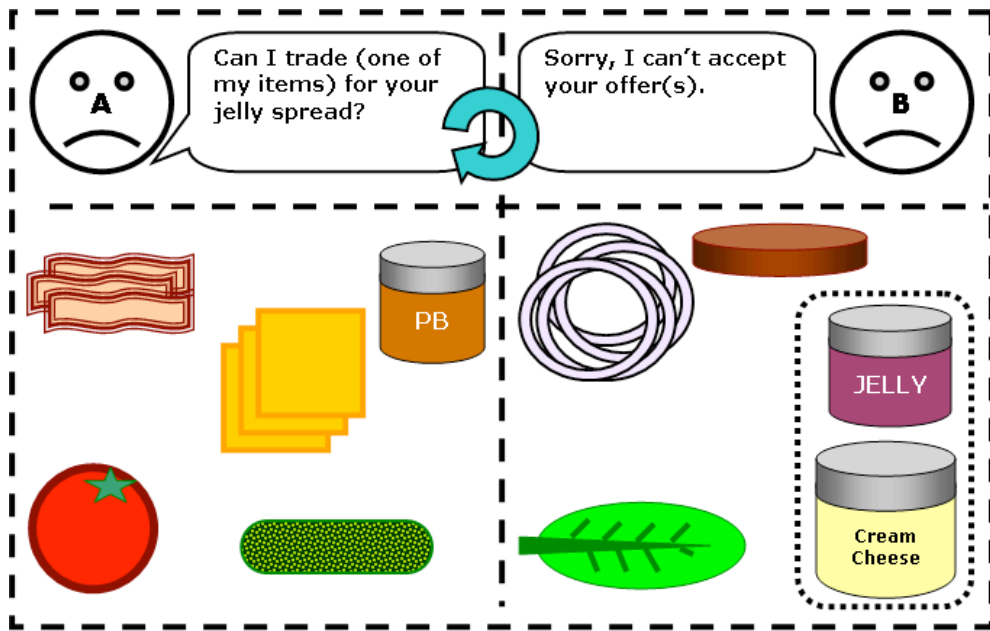
Figure 1.1: Alice can't get Bob to give up his jelly spread
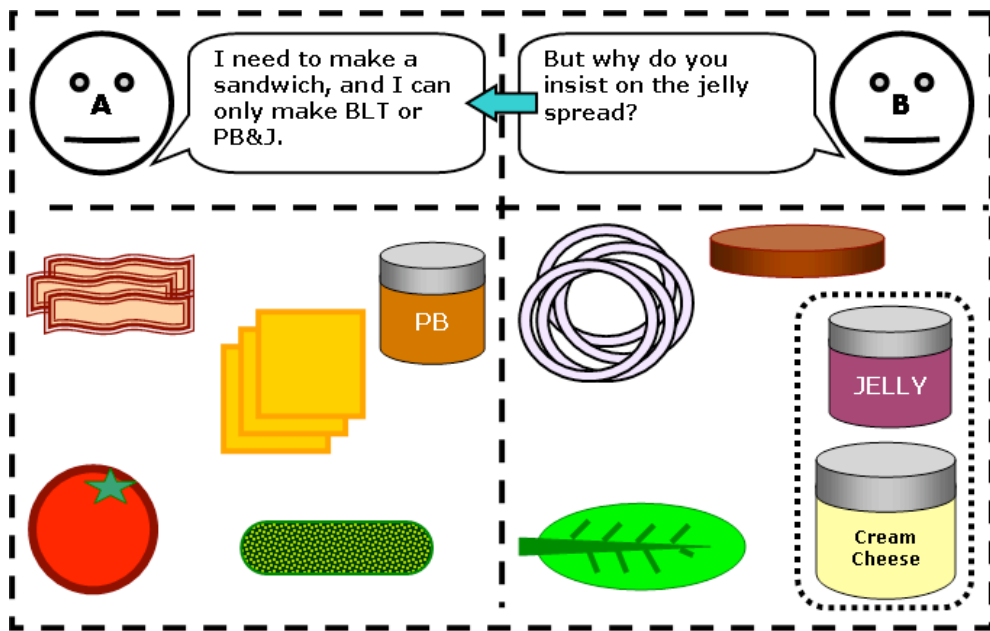
Figure 1.2: Alice reveals her "goal" to Bob

(Figure 1.3), rather than waiting for Alice to figure out that Bob simply will not release his jelly spread to her. In that case, Alice can accomplish her "goal" with the preparation of a BLT, and Bob will still accomplish his with the preparation of the CC&J.

However, if Bob decides to negatively view the fact that Alice's utility from the deal will far outweigh his own, he can additionally ask Alice to throw in the pickle, so that he cam make a cheeseburger later. Going further, he can even demand her peanut butter as well. So he could end up extorting Alice out of her entire surplus of pickle, cheddar and peanut butter, all for the lettuce she desperately needs.

On the other hand, supposing that Alice started out with nothing other than bacon and tomatoes, Bob might sympathetically contribute to the combined welfare of the kitchen and
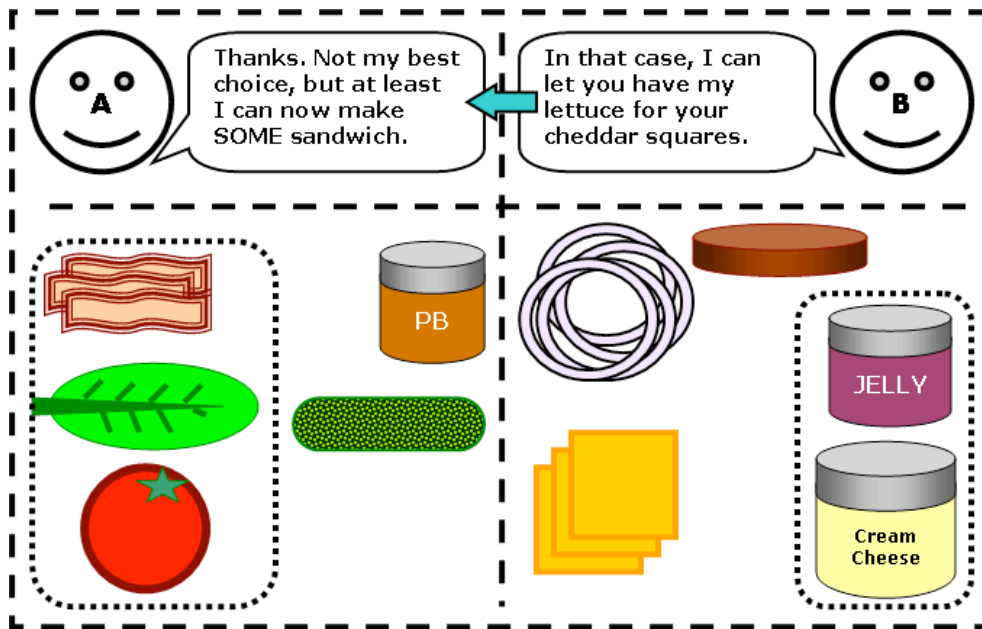
Figure 1.3: Bob makes a simple *quid pro quo* deal with Alice

let her have the lettuce, as he can still reach his goal without it. Alice's communication of her capabilities to Bob, in either case, is analogous to goal revelation in a negotiation scenario, as it exposes the paths that the revealing negotiator needs to traverse, and the resources needed to traverse those paths.

One of the aims of this work is to study situations like this, where the approach to goal revelation by negotiators can be helpful, even sacrificial towards the common good, or apathetic, even greedy in favour of the negotiator's own outcome. The questions on which the dissertation centres are elucidated in the following section.

1.2 Research Questions

The empirical study performed for this dissertation attempts to answer the following questions within the context of goal-oriented resource trading games. The questions centre on experimentation on and analysis of the effects of goal revelation dialogue. This includes the act of asking an opponent to reveal their goal (interest inquiry or solicitation) and the act of revealing one's goal (interest revelation). Another relevant variable is negotiators' status of dependence on or independence of other participants in achieving their primary tasks.

- What kind of platform can effectively enable us to conduct experiments with and investigate aspects of goal revelation in computer-mediated negotiations?

- How does the availability of goal inquiry and revelation options affect negotiation game dynamics in terms of strategy and outcomes? And do any such effects vary due to other features of the negotiation scenario?

1.3 Methodology

Experiments were conducted by having human subjects play a negotiation game over a two negotiation protocols. In the first protocol, negotiators had to make decisions based solely on the knowledge of their own interests, and estimates of their opponents' interests based on

incoming proposals and responses to outgoing proposals. In the second protocol, the negotiators were allowed, in addition, to directly query their opponents for information about the latter's interests, and reveal their own interests when requested to do so. These experiments were run on a modified version of a platform that has a successful history of employment in earlier studies of human and agent negotiation.

Each game was played by two players using a simple visual abstraction of a time-constrained turn-based resource trading board game, with each player having a goal to reach using the available and negotiable resources. The same decision-making settings were played in both protocol experiments, but the sets of players for both experiments were different. Games were set up to encourage negotiation but not to force it, and a decision support tool was provided to aid players in discovering potential alternative paths and optimal offers. Players were given no in-game or game-to-game measures of their performance.

The data from these games was collected and analysed to discern the effects of allowing interest inquiry and revelation on player behaviour and performance outcomes. As both protocol experiments were played using the same decision-making settings, the data was analysed for both general trends across games and any changes in behaviour and performance between protocols over identical decision-making settings.

1.4 Scope

The dissertation attempts to answer the research questions within the scope of a simple, yet possibility-rich analogy of a negotiation game. Due to time and volunteer constraints, only two specific negotiation protocols were experimented with. Only turn-based protocols were implemented for experimentation. While the players' utility function was designed to account for as many real-life utility analogies as possible, various other interesting utility functions, such as all-or-nothing, were not experimented with.

Players did not play repeated games in sequence, and were unaware of the identities of their opponents, making a study of active reciprocity impossible. The volunteer pool was also not common between the position- and interest-based negotiation experiments, so no data could be collected regarding differences in performance of specific players.

1.5 Contribution

The contribution of this dissertation includes the design of an interest-based negotiation protocol for use in strategic settings of incomplete information, and the development of a sound experimental design for studying human performance and human behaviour in this protocol compared to that observed in traditional position-based protocols. This work is the first study that investigates the effects of goal revelation in task settings, and that specifically studies human negotiation in this regard. The suitability of the Colored Trails platform for interest-based negotiation experiments was demonstrated through the adaptation of the same platform for experiments with human negotiators over the two protocols.

The contribution of this dissertation also includes the analysis of the utilization of goal revelation options by human players. This extends to the analysis of how utilization of these options affects the likelihood and extent of negotiators reaching individually and socially beneficial outcomes. Effects on the same of the dependence or independence of human negotiators on one another in achieving their goals were also studied with respect to goal revelation.

1.6   Significance

This dissertation features the first known experiments involving human negotiators in a computer-mediated negotiation setting that features negotiator-triggered interest inquiry and revelation, in games where the benefit or drawback of goal revelation depends in part on opponents' positions. By studying human behavioural patterns set in this variant of interest-based negotiating protocol and this type of game, the dissertation becomes more relevant to the study of negotiations where the strategic implications of having one's interests known by opponents is not clear-cut in terms of being beneficial or harmful.

The findings of this dissertation and the analyses of these findings will provide insight into human behaviour and performance in mediated negotiation scenarios, and aid in the development of computer agent strategies that can be specifically adapted for dealing with human negotiators where voluntary interest inquiry and revelation is involved. Eventually, these findings could be applied in a number of fields.

- In multi-agent systems (MAS), to develop more efficient and/or competitive human-facing negotiation agents, perhaps even ones that are capable of tricking humans into believing that so are they. Furthermore, learning how humans behave in negotiations with goal revelation can lead to improvements in the MAS performance in highly complex negotiation problems that would be too heavy if attempted in a purely rational manner.

- Decision support systems can benefit greatly from the knowledge of how humans utilize and might benefit from goal revelation in negotiation. This knowledge can be used to build negotiation support systems that aid in auction bidding, legal dispute resolution, diplomatic conflict resolution, sales negotiations, etc.

- Psychology and cognitive science can also be enriched by the understanding of how goal revelation affects human negotiation behaviour and outcomes. Such understanding would be useful, for example, in resolving hostage situations, improving outcomes in interrogations of crime suspects, and calming marital discord. In particular, the experimentation platform developed for this work can also provide cognitive scientists with a new tool to further investigate cognitive aspects of goal inquiry and revelation in negotiation.

1.7   Organization

The remainder of the dissertation document is organized as follows.

Chapter 2 describes in further detail the sciences of negotiation and interest-based negotiation, and the known problems that are being studied in those fields. It lists specific studies that lay the groundwork for the dissertation, from the foundations to the state of the art, and describes how those studies are relevant to the dissertation.

This is followed by a description of the features of the platform upon which said experiments were run in the Chapter 3. Chapter 4 that follows describes how the protocols, players, games and game control were set up for the experiments.

Chapter 5 then summarizes the information that was produced by processing the raw data from the experiment, followed by Chapter 6, which discusses the attempted understanding of the empirical results in light of theoretical background and implications. Finally, Chapter 7 describes the foreseeable avenues of exploration complementing or building upon this dissertation, followed by a brief summary of the dissertation' conclusions in Chapter 8.

Chapter 2

Background

This chapter describes in further detail the subject matter of the dissertation. Section 2.1 describes the field of negotiation in general, followed by Section 2.2, a discussion of the concepts of position- and interest-based negotiation that will be adapted for the experiments to come. Section 2.3 describes prior work in negotiation as studied with respect to multi-agent systems. Finally, Section 2.4 describes human negotiation psychology, and prior work related to human behaviour and performance in computer-mediated negotiations with both humans and agents.

2.1 Negotiation

Negotiation is the process by which two or more parties seek to achieve their interests, given the scarcity or sub-optimal distribution among the negotiating parties of the resources required to achieve said interests. The assumption of a party entering a negotiation is that there are resources distributed among other parties that, if acquired by it, would result in greater utility for itself.

DEFINITION 1. For a particular state of an agent's world, *utility* refers to the measurable value associated with that state by the agent. The higher the utility, the more desirable for the agent is the state with which it is associated.

A successful negotiation terminates in either an agreement among the negotiating parties to redistribute resources among themselves, or an accurate consensus among the negotiating parties that no possible redistribution of resources would be more beneficial than the one in existence prior to the negotiation process. An unsuccessful negotiation terminates in either a collapse of the negotiation process due to time constraints, withdrawal due to a breach in the terms of negotiation, or withdrawal because of an inaccurate estimation of the likelihood of an agreement.

DEFINITION 2. When two or more negotiators have committed to redistributing resources among themselves in a particular allocation, or to not doing so, the decision either way is called an *agreement* or a *deal*.

In the following subsections, specific facets of negotiation theory relevant to this dissertation are discussed. The concept of task dependency status will be described in Subsection 2.1.1, and the various negotiating attitudes and situations that negotiators may experience are explained in Subsection 2.1.2. Descriptions of and references to work in distributive and integrative negotiations follows in Subsection 2.1.3.

2.1.1 Task Dependency

A negotiating participant in a negotiation may have a primary ultimate objective, the utility from the achievement of which would overwhelm any gain or loss due to progress or expenditure involved in achieving that objective. A participant may also have a number of milestone objectives that may or may not be necessary to achieve the primary objective. If a participant can achieve its primary objective with the resources it possesses, it may be described as *task independent*. Else, it is said to be *task dependent*(Talman et al., 2005). If a set of players are all incapable of achieving their respective primary objectives, these players are referred to in this work as being *co-dependent*.

2.1.2 Attitude Towards Opponents

A negotiator may have varying attitudes towards its fellow participants' outcomes. This can be depicted on Zhang et al's dual concern model, show in Figure 2.1. When a participant attaches importance only to its own outcome, its attitude toward negotiation is competitive (self-interested); when a participant attaches the same degree of importance to its own outcome as it does to the outcomes of the other participants, its attitude is cooperative; when the participant attaches more importance to the outcomes of other participants at the expense of its own outcome, its attitude is accommodative; if the participant is indifferent to all outcomes, its attitude is avoidant (Zhang et al., 2002). If it does not fit into one of these extremes, it can be described as sharing or compromising. Of course, this attitude can vary from negotiation to negotiation, based on the task dependency status and available resources of the negotiator, and on the perceived dependency status and requirements of its opponents.

DEFINITION 3. An *outcome* is a possible result of a negotiation from the perspective of a given negotiator or group of negotiators.
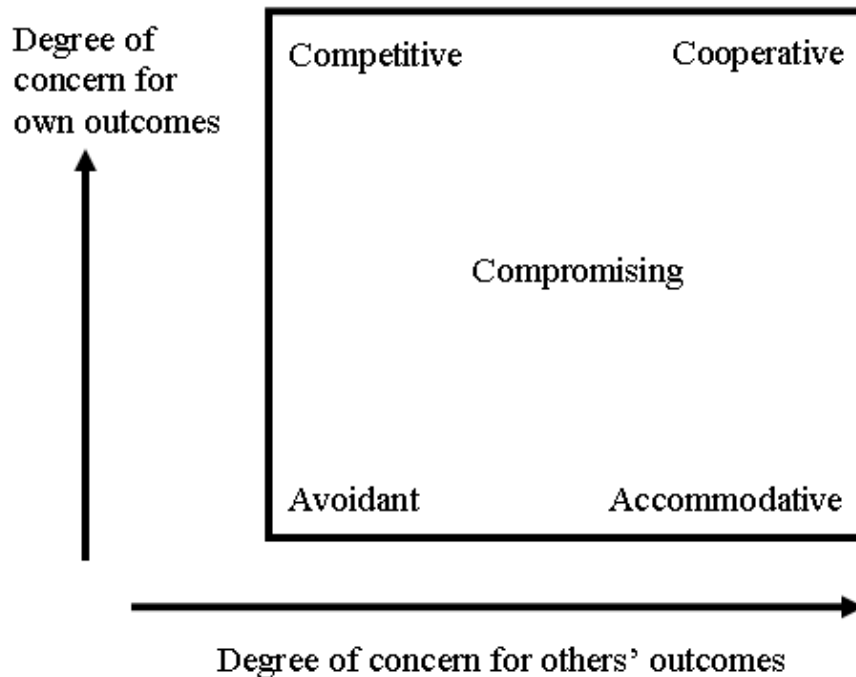

Figure 2.1: The dual concern model of negotiation attitudes

Zlotkin and Rosenschein (1996) lists four possible situations that any two negotiating agents may find themselves in, depending on the game setup and the initial resource allocation.

- A *symmetric cooperative* situation is one in which both agents benefit from each other's existence in the same world, as their cooperation would result in a deal of higher value to both of them, compared to what they would have achieved if each of them had been alone in the game. This would be, in other words, a synergistic situation, where the agents' cooperation would produce a combined reward greater than the sum of that produced by their actions if the other agent did not exist.

- A *symmetric compromise* situation is one where there are individual rational deals for both agents now that they are both in the same world, but where their utility would be greater if they were alone. This would include distributive situations, where the resources of the agents' world have to be divided between them, and there is no way for them to cooperate and increase their individual gains beyond what they would achieve had they been alone in the world.

- A *non-symmetric cooperative/compromise* situation is one in which the best individual rational deal between the two agents would gain more utility for one agent compared to what it would have been able to gain if it were alone in the world, but have the opposite effect on the other agent. Therefore, it is a cooperative arrangement for one of them, but a compromise for the other.

- A *conflict* situation is one in which negotiation is not fruitful due to the lack of either a negotiation set, or of any individual rational deals altogether.

DEFINITION 4. An *individual rational* or *individually rational* deal is a deal that results in a utility for a negotiator that is at least as great as the utility it will get without making any deal.

DEFINITION 5. A deal among agents *Pareto-dominates* another deal if it allows one or more agents to gain utility without forcing one or more other agents to lose utility.

DEFINITION 6. The *negotiation set* is the set of all deals that are individual rational and are not Pareto-dominated by any other deals for the negotiators.

2.1.3 Distributive and Integrative Negotiation

Distributive negotiations are based on the concept of "cutting up a pie," where the size of the pie is fixed. They occur when the problem is, or is perceived as a "zero-sum game," where as the value along a single dimension shifts in either direction, one side is better off and the other is worse off. In such cases, the negotiators focus on mutually exclusive goals (Kersten, 2001), either with an all-or-nothing view of their utility or with the proportion of the contested resources they acquire deemed to be positively correlated to their utility. Examples of these include auctions and airline ticket reservations.

Integrative negotiations, on the other hand, are based on the concept of "expanding the pie." They occur when the participants or mediators find ways to resolve disputes using information about the actual needs and priorities of the participants. Integrative negotiations will naturally be multi-issue negotiations, as single-issue negotiations will force the participants to compete over possession or partitioning of a uniform pool or resources. If the participants have different requirements and priorities in reaching their respective goals, there exists a possibility that the existing resources can be distributed in a way that each participant can still do so. In such multi-issue negotiations, the opportunity for joint improvement is provided by the difference in priority and weight attached to the different issues by individual agents (Coehoorn and Jennings, 2004).

2.2 Position-based and Interest-Based Negotiation

A *negotiation position* of an agent can be defined in terms of the resource(s) that agent wants to acquire from other negotiators. Position-based negotiation, or positional bargaining, oc-

curs one where the participants consider only their own negotiation positions, and where the dialogue between participants is focused on the same. Examples of positional negotiations include alternating-offer bargaining and auctions Rahwan et al. (2003). Any change in position is brought about solely by lack of progress in reaching an agreement with the position already held.

In interest-based negotiation, on the other hand, agents may argue about each other's beliefs and other mental attitudes in order to justify their negotiation positions, and attempt to influence each other's negotiation positions. By understanding the reasons behind positions, a problem can be redefined in terms of the underlying interests. By discussing and considering these interests, participants are more likely to reach a mutually acceptable agreement (Rahwan et al., 2003).

Therefore, while a multi-issue position-based negotiation can involve attempts at distributive bargaining progressing to integrative bargaining through trial-and-error with incomplete information, a multi-issue interest-based negotiation becomes integrative once the participants' interests are made known to their fellow participants, followed by a distributive bargaining phase to split up the value addition.

2.3 Negotiation and Multi-Agent Systems

A multi-agent system is a system of loosely-coupled interactive intelligent computer agents. Work on negotiation in multi-agent systems is rich with both theoretical and empirical studies, some of which follow in this section.

Rubinstein (1985) provided a theoretical model for prescribing negotiating strategies in settings of repeated interactions that are optimal under certain conditions (e.g., participants are rational and consistent in their beliefs about each others' objectives). Zlotkin and Rosenschein (1996) explored the concept of state-oriented domains, where the description of the world is modelled as a state, and where operators cause the world to move from one state to another. It also proposed multi-agent Unified Negotiation Protocols - a set of product-maximizing mechanisms based on either semi-cooperative deals or multi-plan deals, aiding in both conflict resolution and cooperative agreement. Faratin et al. (2002) developed an algorithm for the negotiation of multi-feature contracts using similarity heuristics over iterative offers, proposals and counter-proposals to vary agents' own offers towards their opponent's offers without losing value for themselves.

DEFINITION 7. An *offer* or *proposal* is a possible reallocation of resources to which the proposing agent is willing to commit, but only if the other agents involved in the reallocation also commit to it.

Some "noisy" estimation methods like kernel density estimation (Coehoorn and Jennings, 2004) or Bayesian refinement (Buffett and Spencer, 2005; Hindriks and Tykhonov, 2008) have been successfully developed and demonstrated for use in bilateral position-based negotiation to help agents figure out what their opponent needs. Winoto et al. (2004) used formal analysis and simulations to show that use of non-monotonic-offers protocols leads to higher surpluses and fewer aborted negotiations compared to monotonic-offers protocols, as long as the negotiations are not dominated by agents who make arbitrary or bullying offers, or who only accept monotonic bargaining.

DEFINITION 8. With each sequential offer in *monotonic* bargaining, agents are only allowed to raise or maintain the utility of the offer to their opponents compared to that of the previous offer.

Soo and Hung (2002) performed experiments with cooperative negotiators in games with a limited number of messages, and implemented agents that could incrementally learn from each of the other agent's proposals. They demonstrated empirically that such negotiation learning agents could get closer to the Pareto efficiency agreement much faster than such non-learning

negotiating agents as simple random agents, rational agents, and cooperative agents. Narayanan and Jennings (2005) defined an adaptive negotiation model and cast it as a Markov Decision Process, using a multi-agent negotiation simulation to demonstrate its ability to train optimal negotiation policies in a dynamic e-commerce environment. Lai et al. (2007) presented a model for alternating offer multi-issue negotiations for agents with non-linear utility functions, employing a moving baseline to make the negotiation space search tractable, and using a simulation to empirically demonstrate that the model asymptotically approaches Pareto optimality.

DEFINITION 9. An agreement is *Pareto efficient/optimal* if there is no other agreement where one agent can gain additional utility at the expense of any other agents' utilities (that is, no other agreement Pareto-dominates it).

Some studies have also been conducted on the topic of combinatorial auctions (Cramton et al., 2004), where negotiators can bid for packages of items together. Koenig et al. (2006) empirically showed that teams of agents playing using a coordination system based on sequential single-item auctions performs almost as good as would be expected from an optimal combinatorial auctions and much better than a coordination system based on parallel single-item auctions. Archer et al. (2003) discovered methods to make combinatorial auctions truthful by creating an incentive compatible version of the randomized-rounding approximation algorithm, resulting in an approximation algorithm that makes honesty in the agents' best interests.

DEFINITION 10. An *incentive compatible* design is one that makes honest reporting of private information part of the dominant strategy of all negotiating agents (Babaioff and Walsh, 2003).

Talman et al. (2005) describes a model where agents have to exchange resources to achieve their goals, while not knowing each other's available resources. It describes how helpful (cooperative and reliable) agents that are task independent engage in benevolent behaviour and are taken advantage by unhelpful agents. The source also notes, though, that task-independent helpful players do better than unhelpful players overall, as the latter do not realize the full potential of cooperating with others. This dissertation will analyse human negotiators' performance in an mediated interest-based negotiation setting where negotiators are aware of each others' resources, and study the relationship between task dependency, interest revelation, performance and "helpfulness" for human negotiators.

Work in automated negotiation in Artificial Intelligence (AI) has proposed algorithms for argumentative strategies which support or attack the different positions of parties in a negotiation (Kraus et al., 1998; Rahwan et al., 2003). An argumentation-based negotiation protocol that links offers and arguments was proposed by Kakas and Moraitis (2006), describing the extension of negotiation strategies within other multi-agent negotiation mechanisms, and proving the advantages such an approach offers in allowing agents to dynamically adopt strategies in response to their environments. Argumentation-based negotiation algorithms have been used by computational agents and several works have studied conditions under which such strategies outperform position-based protocols (Pasquier et al., 2007; Rahwan et al., 2007). This dissertation directly extends these studies by showing that argumentative-type protocols are advantageous to people, and by introducing a form of "argumentation by proxy" through the use of interest inquiry and revelation dialogue.

DEFINITION 11. *Argumentation* is defined as the construction of statements of position or support, and logical evaluation of those statements in relation to other such statements.

2.4   Negotiation and Humans

Studying the behaviour of human negotiators is complicated by various sub-optimal and non-rational actions that are observed when humans negotiate. These can be due to emotional or cultural factors. Sometimes, these behaviours are consistent with the known phenomenon of "bounded rationality", which is observed in human reasoning and decision-making (Simon,

1957). Some of these observed behaviours are summarized in Hindriks and Jonker (2008), and are listed and described here.

- Making too large concessions, thereby proposing for or accepting too-small a share of the bargaining pie in a distributive situation or phase. This may happen when a person is too afraid of rejection, and offers their entire surplus.

- Failing to recognize and exploit the opportunity to create value in a potentially integrative situation.

- Rejecting an offer that is better than any other available option. This could happen as a result of suspicion or as a punitive action.

- Settling for terms worse than alternative options, such as when humans agree to an offer even when it is clearly worse than other known alternatives, perhaps out of a sense of sympathy or guilt.

Hindriks and Jonker (2008) also lists some causes for such behaviour.

- Lack of training, or untimely, vague and inaccurate feedback resulting in self-reinforcing incompetence.

- Lack of preparation, leaving the negotiator unaware of an important part of the bargaining pie and/or the preferences and circumstances of the parties involved

- Structural barriers to agreement, including die-hard bargainers, cultural and gender differences, disruptive or in communicative opponents, and communication failures

- Mental errors, such as the escalation error, biased perception, irrational expectations, overconfidence, and unchecked emotions.

- *Satisficing* (Simon, 1957), the practice of not seeking to maximize outcomes, but settling for "good enough", due to uncertainty of the future, the costs of acquiring information, and the limitations of human computational capacities. This forces the humans to act within "bounded rationality".

DEFINITION 12. *Bounded rationality* is a decision-making process whereby the cost of time, computational power and information management lead an agent to settle for an outcome that falls short of an outcome of superior utility that would have been reached if the aforementioned costs were not considered.

Culture may also influence human negotiation behaviour. This influence has been modelled along five dimensions of cultural characteristics in the *Hofstede model*, described in Hofstede et al. (2008).

- Affiliation (collectivism vs. individualism): Measured between complete group welfare concern and total individual welfare concern.

- Hierarchy (large vs. small power distance): Measured between between highly stratified authoritative hierarchy and more level egalitarian power-sharing.

- Aggression and Gender (masculinity vs. femininity): Measured between a highly masculine heavy-handed fighting society, and a strongly peace-loving, consensus-seeking society.

- Otherness and Truth (strict uncertainty avoidance vs exploratory tolerance): Measured between a society of strict rules, rituals and absolutism, and one that encourages curiosity, exploration, diversity and novelty.

- Urgency of Gratification (short- vs. long-term): Measured between attitudes of "living for today" with conspicuous consumption and strict norms of honour and propriety, and of "saving for a rainy day" with pragmatism, planning, perseverance and calculated foresight.

Hofstede et al. (2008) modelled these dimensions using the ABMP bargaining model (which has a utility function based on expected gain, quality and risk) and concluded that this enhanced model can be used to develop agents for cross-cultural bargaining. Further on the topic of gender, Katz and Kraus (2007) performed empirical evidence that using different strategies of negotiation against male and female negotiators improved performance, and presented a learning algorithm that could provide an online indicator of the gender separability-level of the population which further increased performance. Weingart et al. (1996) showed that human players trained in specific negotiation tactics before negotiating were able to perform better in integrative negotiations, but not in distributive negotiations.

DEFINITION 13. A *utility function* of an agent is a function that takes the state of the world as an input and returns the utility as output.

There are few works offering an empirical analysis of peoples negotiation strategies in repeated interactions. Work in the psychological literature about strategic interaction has focused on specific domains (e.g., seller-buyer disputes (Loewenstein et al., 1989), Middle East peace talks (Atran et al., 2007) or completely abstract settings like the prisoners' dilemma. Loewenstein and Brett (2007) conducted a study which looked at how goal framing prior to the negotiation procedure affects strategy revision. None of these works have compared the effects of goal revelation directly within repeated negotiation.

DEFINITION 14. *Goal framing* refers to how an agent represents its goals in terms of specific sub-goals, interests or positions.

Bosse and Jonker (2005) presented two experiments involving humans and rational negotiation software agents on a platform different from the one used in this dissertation. The results indicate that the negotiations are fairer when conducted with their agents alone, compared to negotiations that pitted humans against agents, and that when pitted against agents, humans did better than agents in terms of individual performance. An experiment featuring humans against agents was also conducted in Lin et al. (2008), where the agents were playing using a bounded rationality model, and the negotiators had to do with incomplete information. In this case, the agents did better than the humans, achieving more and individually better agreements. However, these were performed using position-based negotiation games, not interest-based games like the ones used in this dissertation. As was the human negotiation experiment by Bosse et al. (2004) performed on the same platform as that used in Bosse and Jonker (2005), showing that humans in multi-issue negotiation games find it difficult to guess where the Pareto Efficient Frontier is located, making it difficult for them to accept a proposal.

Work by Heiskanen et al. (2001) has found that people minimize the amount of private information they reveal in negotiation to avoid weakening their positions, while Vorauer and Claude (n.d.) observes that human negotiators tend to overestimate the accuracy of their opponents' goal estimations the more they know about their opponents' goals and the less constrained their communication protocol is. This dissertation experiments over two levels of communication constraint severity, and shows that under certain conditions people are willing to disclose private information to others, which results in more efficient agreements.

Chapter 3

Colored Trails Platform

Colored Trails (CT) is a testbed developed for the purpose of investigating the decision-making that arises in task settings, where the key interactions are among goals, tasks required to accomplish those goals, and resources needed to perform the tasks (Grosz et al., 2004). The experiments performed for the purpose of producing this dissertation's results used special-purpose adaptations of CT's third version to permit the negotiation protocols of interest to be implemented in the gameplay.[1]

The elements and gameplay of a Colored Trails negotiation game are described in Section 3.1 and Section 3.2 respectively. The various panels and actions in the modified interface are described next in Section 3.3, followed by the analogy between a game on the platform and a real-life negotiation scenario in Section 3.4.

3.1   Game Elements

The main elements in a CT game are the board, the players, the goals. the chips and the scoring function.

The board is a rectangular coloured tile array, the size, colours and layout of which can be specified by a dimensional, palette and grid specifications. The players are represented on the board by custom icons, and can be uniquely identified by a unique in-game ID, as well as a pin number across games. Their actions can be controlled either through a graphical user interface by a human player or by a software agent.

Goals are represented by custom icons as well, and can either be designated for specific players or open to all players. The chips serve as resources for trading, coloured according to a specific palette and allocated to each player by the game setup. Generally, the chips will be coloured using the same palette as that which was used to colour the game board.

The scoring function allows the game creator to set the rewards or penalties for reaching the goal, for each chip the player has left at the end of a game, and for each unit of distance the player the player remains from its goal after attempting to move towards it. In order to introduce incentives against lengthy negotiation, the scoring function was extended to include a cost for the number of dialogue rounds the players went through before their game ended.

All of these elements are packaged together into a single game state object, which can be persisted for repeat use.

---

[1]Downloads of and more information about Colored Trails are available at the project website, http://viki.eecs.harvard.edu/confluence/display/coloredtrailshome/Home

3.2 Gameplay

A CT game is centred around a running server application, which hosts the players and runs the game configuration. The players connect to the server and wait for the controller to set them up with a game and opponent(s). Various details of the game, such as the positions and goals of other players may be hidden from and revealed to them by the game configuration, which also controls the players' roles, turns, availability of actions and phase transitions. A modification was made to the player states to control whether a player's goals were visible to its opponents.

Players are assigned an initial position on the board, and a set of chips. The objective of a player is to move towards the goals that are open to it i.e. either goals specific to the player or generic goals open to all players. In order to move to an adjacent tile, a player needs to give up a chip that is of the same colour as that tile. Chips may be transferred among players through negotiations wherein a player offers certain chips from her own chipset in exchange for certain chips from her opponent's chipset. Players cannot ask for chips their opponents do not have, and cannot offer chips they do not have. Negotiation, exchange, movement and scoring take place over three phases.

The first phase is the communication phase, during which players may exchange negotiation dialogue messages. At the end of the communication phase, the movement phase starts with the execution of any exchange that may have been agreed upon by the players, followed by the server moving each player so that its score is maximized given its post-exchange chips and goals. After they have moved, the scores are computed for each player. Finally, the feedback phase allows the players to see the results of their movements. Each phase can last for a certain amount of time, which is also configurable.

3.3 Game Interface

The user interface consists of eight panels which are displayed by the controller as determined by the gameplay configuration. (figure for each)

- The taskbar contains the main action buttons for the players, allowing them to connect to a server and propose a transaction. A button for goal revelation was added to this panel for the IBN games. Figure 3.1 depicts an example of a taskbar in an IBN game right after a goal inquiry from the responder has been received, permitting the proposer to either make a new proposal or reveal her goal.
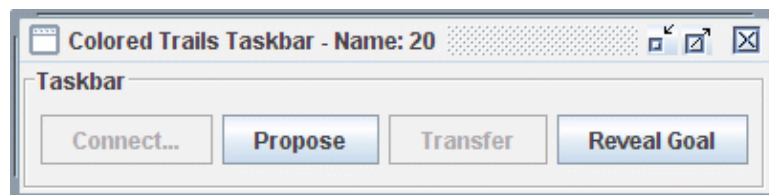
Figure 3.1: CT3 GUI Taskbar in an IBN Game

- The main panel displays the game board and the chipsets of the players (filtered appropriately), also displaying the current phase and how much time remains before it ends. Figure 3.2 depicts an example of a main panel during the Communication Phase, where the player's opponent's goal is not visible and the player has selected a particular path to highlight on the board.

- The proposal composition window allows players to select the number of chips of each colour that they wish to offer and request, and to send this proposal to their opponents. It does not allow players to offer more chips than they have, or request more chips than

Figure 3.2: CT3 GUI Main Panel

are in the possession of the opponent to whom they are proposing. Figure 3.3 depicts an example of a proposal composition window as seen when a player is preparing to request a green chip in exchange for a red chip.

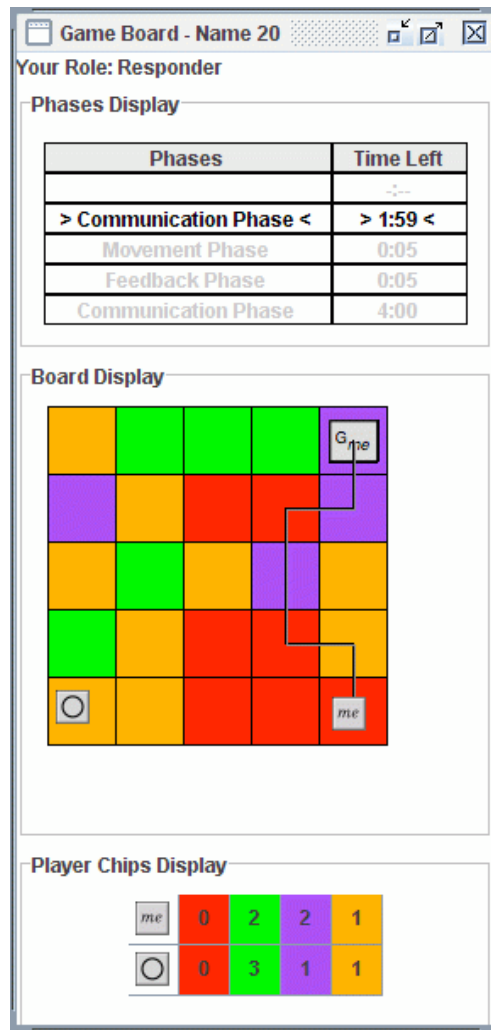- The response action window allows a player to whom a proposal has been sent to reply with rejection or acceptance. After the response has been sent, the window displays the response that was sent. This window was extended with a button that allows responders to request the proposer to reveal her goal. Figure 3.4 depicts an example of a response action window in an IBN game right after a proposal has been received, allowing the responder one of the three available options by clicking one of the three enabled buttons.

- The proposal tracking window displays the sent offer, waiting for the response to the proposal and displaying it once it has been received.

- The response tracking window appears after a response has been sent, displaying the response message.

- The pathfinder window is a decision support tool that shows the players what paths they can take to reach their goals, as well as information indicating how beneficial each path would be. The pathfinder also allows players to view the path options for their opponents, if the game state permits. A chip count changer set into the pathfinder lets players see the hypothetical changes to their path options if they were to add or subtract chips of
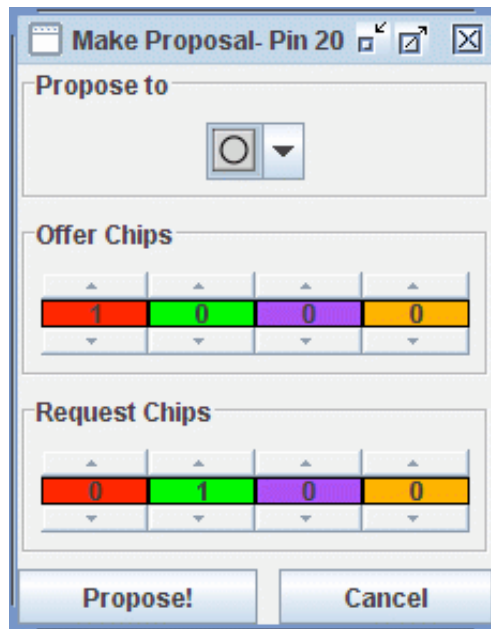
Figure 3.3: CT3 GUI Proposal Composition Window
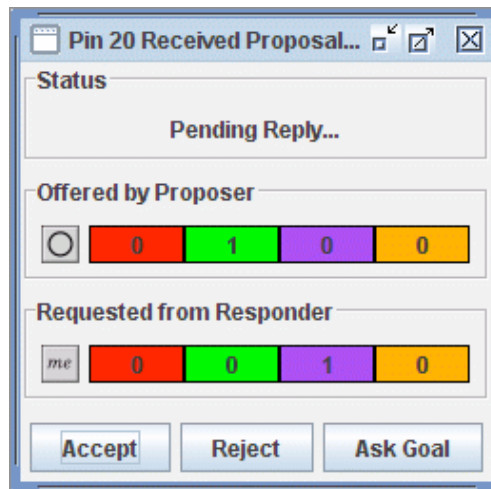
Figure 3.4: CT3 GUI Response Action Window in an IBN Game

each colour. In order to facilitate the discovery of alternative paths and to make potential exchanges easier to visualize, the pathfinder was modified to show not only the chips the player would need to get to his goal, but also the chips the player would have in surplus if he got to his goal, the number of tiles remaining between the player and the goal if he made his best move, and the number of tiles he would be able to traverse in his best move along the path. Figure 3.5 depicts an example of a pathfinder window where the shortest path the player's goal has been highlighted for viewing on the game board in the main panel.

• The mood window appears at the start of each game, forcing the player to engage in a neutral activity that is designed to minimize the effects of past games on the player's behaviour in the game to come. In this experiment, it asks the player to choose from three emotion icons one that best represents their mood.
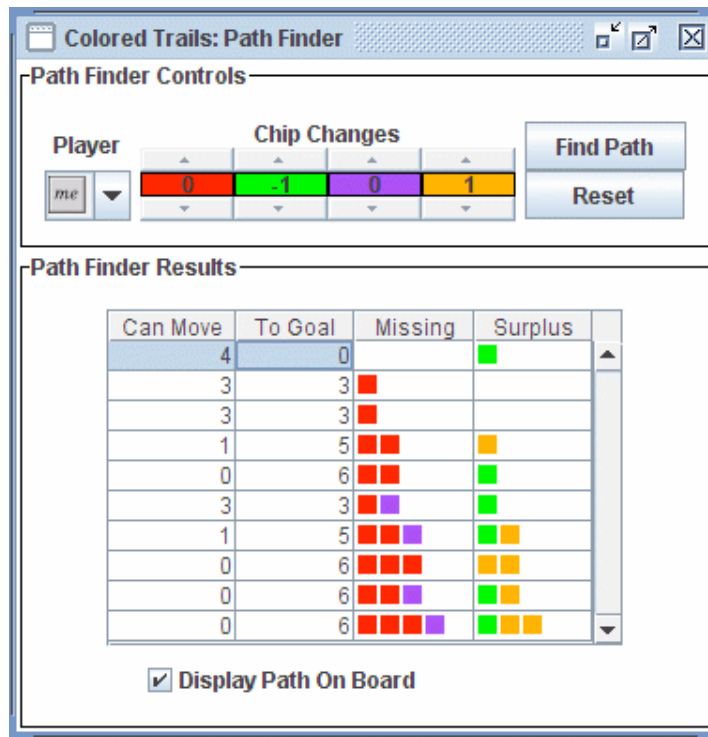
Figure 3.5: CT3 GUI PathFinder Window

3.4 Task Analogy

CT provides a realistic analogy to task settings, highlighting the interaction among goals, tasks required to achieve these goals, and resources needed for completing tasks. Chips correspond to agent resources and skills required to fulfil tasks. Different tiles on the board represent different tasks with different requirements. A players possession of a chip of a certain colour corresponds to having the required ability available for use when needed. Not all players possess chips in all colours, much as different agents vary in their capabilities. Traversing a path through the board corresponds to performing a complex task whose constituents are the individual tasks represented by the colours of each tile. It has been shown that people are more likely to engage in cooperative behaviour when using this game than when using completely abstract representations such as payoff matrices (Gal et al., 2007).

Chapter 4

Experimental Setup

This chapter describes and justifies the decisions that were taken in the setup of the experiment, specifically with respect to the design of the protocols, the configuration of the games, the design of the player match controller, and the preparation of the players.

The specific variants of the position- and interest-based negotiation protocols that were used for the experiments are described in Section 4.1. This is followed by a description of the setup of the experiment's game configurations in Section 4.2 and the controller design in Section 4.3. Finally, the selection and preparation of the participants is described in Section 4.4.

4.1 Negotiation Protocols

Two negotiation protocols were employed in the empirical study in order to isolate and analyse the effects of interest information dialogue on negotiation.

DEFINITION 15. A *negotiation protocol* is a definition of the types of dialogues that can be used in the negotiation, what content they may contain, when they may be used, and by which negotiator(s) they may be used.

The first was a position-based negotiation (PBN) protocol wherein the agents had to make their proposals based on information limited to the positions and resource pools of their opponents, and attempt to guess their opponents' resource deficits and surpluses based on their opponents' proposals and their opponents' reactions to their proposals. This protocol establishes a baseline for agent performance and behaviour. This protocol is described in Subsection 4.1.1.

The second was an interest-based negotiation (IBN) protocol that extends the position-based negotiation protocol by allowing agents to respond to a proposal by inquiring into their opponents' goals, and by allowing agents to reveal their goals to their opponents should they be asked to do so. This protocol is described in Subsection 4.1.2.

4.1.1 Position-Based Negotiation

The PBN protocol is based on the alternating offer protocol. The protocol operates in a turn-based mode, with one agent starting out in the role of the proposer and the other agent starting out in the role of a responder. The initial role allocation is random. Both agents can only see their opponents' positions and resources. The proposer may choose resources from her resource pool to offer her opponent, choose resources from her opponent's resource pool to request in return, and propose this transaction to the responder.

The responder can accept the proposal, in which case the agents' resources are allocated

accordingly and the negotiation terminates. Alternatively, the responder can reject the proposal, in which case the negotiation continues with the responder given the role of proposer and vice-versa. The responder or proposer can also opt out of the negotiation during their turn, in which case the negotiation terminates in a conflict allocation, i.e. no change in the resource pools of the agents. Once a proposal has been rejected by a responder, it is considered dismissed, and the responder may not retract his rejection; he must either propose the same transaction in reverse, or wait until his opponent makes the same proposal again. The state diagram for this protocol is shown in Figure 4.1.
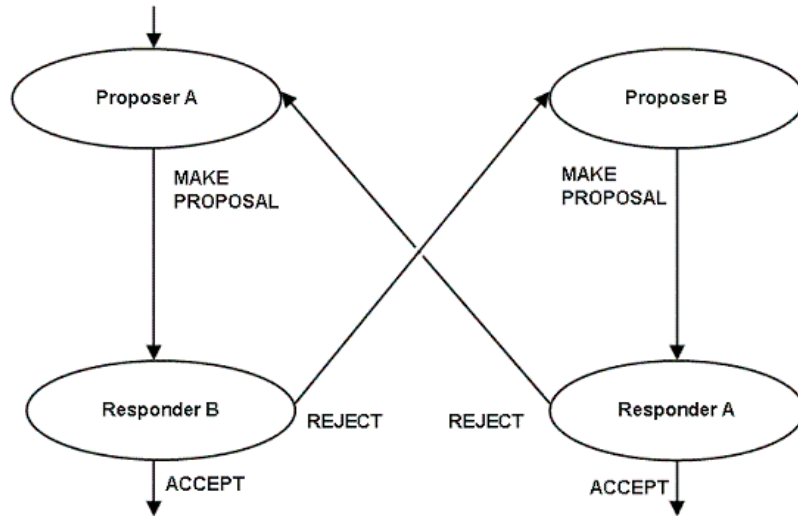

Figure 4.1: State Flow of Position-Based Negotiation Protocol

4.1.2   Interest-Based Negotiation

The IBN protocol extends the PBN protocol by allowing an additional response to a proposal: in addition to opting out, accepting or rejecting a proposal, a responder may send a message back to the proposer asking the proposer to reveal her goal to the responder. An agent can only ask for his opponent's goal when he is in the role of a responder, and only if he has not already revealed his goal.

In this case, the proposer can either opt out, make a new proposal, or reveal her goal to the responder. Should the proposer reveal her goal, the responder and proposer switch roles and the game continues in the turn of the new proposer. An agent can only reveal her goal to her opponent when she has been asked to do so. Should she choose to make a new proposal instead of then revealing her goal, she must wait to be asked for her goal again in order to reveal it. The state diagram for this protocol is shown in Figure 4.2.

4.2   Game Setup

The games played in the experiments are set up to facilitate a large space of possible transactions, allow multiple paths of movement towards the goals, force at least one of the players to negotiate, discourage the players from dragging out negotiations, and encourage the players to be conservative with their resources.

In order to keep the game simple for the players, the game board is set up as a 5-by-5 tile layout, with a palette of 4 colors for board tiles and player chips. Two players start out in random locations, and each has a randomly-located goal, visible at first to them alone. Each
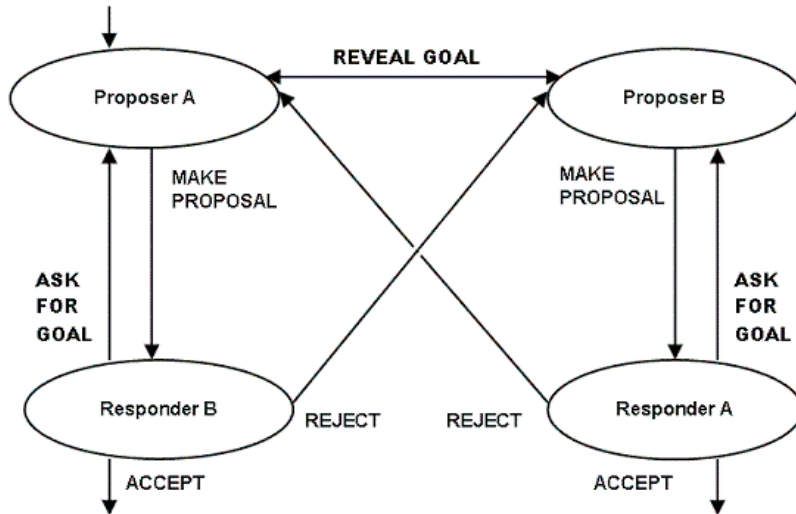
Figure 4.2: State Flow of Interest-Based Negotiation Protocol

player is initially allotted 7 random chips, to increase the possibility of there being multiple possible paths to their goals.

A number of constraints apply to the setup of the games. Both players must not be able to reach their respective goals on their own, ensuring that at least one player is forced to attempt negotiating a transaction in order to reach its goal. There must be some possible combination in the union of resources allotted to both players which permits at least one player to reach its goal. Either player can reach its goal with the other's help, but it may not be possible for both players to reach the goal. This ensures that in any game, it is potentially possible for at least one player to reach its goal. As a result, the game setup is designed to produce a bilateral negotiating situation that can be symmetric cooperative.

The positions of both players and both goals must never share the same tile on the board, which makes the board layout clearer to GUI players by preventing different icons from over-lapping, and prevents players from being placed on top of their own goals. Each player and her goal must be in either the same row or column of the board, allowing the player only one shortest path to the goal, so that she may be encouraged consider other paths.

The scoring function used for both PBN and IBN games gives the highest weightage of 100 points for reaching the goal, upon which it adds 10 points for each chip left over after moving. There is also a penalty of 15 points for each minimal Manhattan distance unit the players remain from the goal after moving, and 5 points for each rejected offer of transaction. Rejected offers counted for this purpose include those that have been met with goal revelation requests, but do not include rejections of blank offers, and goal inquiries or revelations. The intended net effect is to encourage players to reach their goals first and foremost, move as close to their goals as possible otherwise, trade away and use as few chips as possible, and reach agreements with as few transaction offers as possible.

4.3   Controller Design

A CT game controller was adapted for the PBN and IBN protocols. In both cases, the game controller allows the players to play in turns, and disables the players' action buttons if it is the other player's turn to send or respond to messages. The players are allowed a maximum of 4 minutes of communication phase to exchange offers or information until they reach agreement. If the players do not reach an agreement within 4 minutes, the communication phase terminates with no transaction assumed, a condition known as the *no negotiation alternative* or NNA. If

they reach an agreement, the transaction thereby agreed upon is registered and the game transitions to the movement phase.

A game generator utility was developed to compose and load prefabricated game pools according to the requires specification and constraints, and an experiment controller was developed to pair players and assign them a number of games from the pool as required. The reason for prefabricating and persisting games into a pool is that the game pool can be shared between the PBN and IBN runs of the experiment, and the player performance in both protocols can be compared for identical decision-making settings or the same set of games.

The experiment controller detects the players logged onto the game host server and loads a pool of games. It pairs up available players and sequentially distributes one game from the pool to each pair, initializing it with the configuration class corresponding to the experiment's negotiation protocol. When all the games allotted this way have terminated, the controller resumes pairing available players and allotting games. It continues doing this until each player has played every other player, once beginning as a proposer and once more beginning as a responder.

For each game, CT's logger was used to persist the initial game state, record proposals and other messages, and log NNA scores, proposal-hypothetical scores and final scores for each player.

4.4   Participants

Twenty-two subjects participated in the experiment, drawn from a pool of students and adults residing in the Boston area. Twelve people participated in the IBN condition while ten people participated in the PBN condition. Each person was given an identical 30 minute tutorial on CT, and played three practice rounds. Each subject was identified by a serial number, was seated in front of a terminal for the entire length of the experiment, and could not see or speak to any of the other participants.

Subjects were not informed of the identities of their opponents in the games they played, either onscreen or by the supervisor. They were paid in a manner consistent with their aggregate scores in all of the games they played. In addition, subjects scores were not revealed at any point during the experiment. Between each game, players engaged in a neutral activity which did not affect their payment (answering questions about their mood), designed to minimize the effects of past games on their future performance.

Chapter 5

Results and Analysis

65 uniquely-configured games were eventually played in both conditions - some more than once - among the human players for each condition. In 14 of these games, players were co-dependent on each other to get to their respective goals, and in the other 51 games, only one of the players was independent of the other player as far as reaching their goals was concerned. Players that queried their opponents about the latter's goal are henceforth referred to as "goal solicitors," and those players that subsequently revealed their goal will be referred to as "goal revealers". Despite there being no interest-related dialogue in the PBN condition, the term "goal revealers" is used therein to refer to those players in the PBN condition that align with the goal revealers in the same settings as in the IBN condition. The same applies the use of the term "goal solicitors" in PBN games).

The following sections describe specific analyses of the results from the experiments. The analysis across different variables of task dependency, protocol, and IBN dialogue utilization is described in terms of the performance of the players in Section 5.1, and in terms of the players' likelihood of agreement in Section 5.2. Finally, the players' behaviour in terms of goal inquiry and revelation dialogue utilization is presented in Section 5.3, and various observations about the patterns seen in the analysis are presented in Section 5.4.

5.1 Performance

This section analyses the effect of goal revelation and task dependency on individual and social outcomes.

	Games with Goal Revelation(s)	All Games
Revealing Player	**41**/23	38/38
Soliciting Player	7/7	34/32
Combined	**49**/32	**72**/61

Table 5.1: Average benefit for players in IBN/PBN conditions for different number of goal revelations (significant difference in bold)

Table 5.1 shows the average benefit to revealers and solicitors in games played in both the IBN condition and the PBN condition (left and right entries within the same cells). The benefit to a player in a game is defined as the difference between the proposed outcome score and the no negotiation alternative (NNA) score. Therefore, if no agreement is reached, a player's benefit is zero. The results are measured with respect to the total set of games played in both conditions (column marked "All Games") and with respect to the games in which one or both players revealed their goals in the IBN condition (column marked "Goal Revelations"). In the latter case, the figures for the PBN side come from the scores of the players whose game roles match

the roles of the revealers and solicitors for the same games in the IBN protocol.

As shown in the table, the combined average benefit for players in the IBN condition was 49, which is 17 points greater than the average PBN benefit in cases where there was at least one goal revelation (SE = 5, paired t-test $t(29)$=1.7, p=0.04). There was an insignificant 5-point difference between conditions in the combined benefit for those games in which no revelations occurred ($SE = 2$, paired t-test $t(34) = 0.5$; $p = 0.28$). This is because there was a combined advantage to players in the IBN condition when the number of revelations was greater than zero, but not when this number equalled zero. In addition, it turns out that this advantage is significant enough to affect players' total performance. Also, as shown by the table, the combined average benefit for both classes of player in all games was 72 in the IBN condition, which is 11 points greater than in the PBN condition ($SE = 3$, paired test $t(64) = 1.60$, $p = 0.04$). In total, there were more goal solicitors than goal revealers, and thus the numbers in the "combined" row do not necessarily equal the summation of "revealing" and "soliciting" players. Therefore, not only does revelation itself have a positive effect on players' social benefit, merely playing the IBN protocol improves the social benefit of both players compared to playing the PBN protocol.

Table 5.1 shows that in both conditions, the relative benefit to goal solicitors from playing the IBN protocol instead of the PBN protocol is significantly less than the relative benefit to goal revealers. Specifically, revealers in the IBN condition increased their average benefit by 18 points, as compared to their performance in the corresponding games in the PBN condition ($SE = 3$, $t(29) = 1.7$, $p = 0.04$). However, there was no significant difference in the benefit between those players that solicited goals in the IBN condition and the corresponding games in the PBN condition (both 7 points).

	Games with No Goal Revelation	Games with Goal Revelation(s)	All Games
Independent Player	**19**/0	-9/-15	**15**/-2
Dependent Player	49/59	**56**/35	40/50

Table 5.2: Average benefit for players in IBN/PBN conditions for different player dependencies (significant difference in bold)

Table 5.2 shows the benefit to players as a function of their dependencies, as well as the number of goals they revealed. As shown by the table, the benefit for dependent players was consistently higher than the benefit for independent players, across conditions. In particular, in those games in the IBN condition where at least one goal was revealed (and the corresponding games in the PBN conditions) they willing to make a sacrifice. This generosity is considerably more effective in the IBN condition, where the average benefit to dependent players is 19 points greater than in the PBN condition ($SE = 2.3$, $t(26) = 2$, $p = 0.02$). Despite the sacrifice incurred by independent players, their overall benefit in the IBN condition is 15 points, 17 points greater than the benefit of players in the PBN condition (SE = 2.3, t(48) = 2.3, p = 0.01). Lastly, we note that there were significantly more dependent revealers than independent revealers (27 vs. 8), and more independent solicitors than dependent solicitors (24 vs. 11). This explains the finding of Table 5.1 that the benefit of solicitors is significantly less than the benefit of revealers. The combined performance for co-dependent games was also found to be higher than that of single-dependent games, but this difference was not significant, perhaps because of the low number of co-dependent games.

Considering all the games played in IBN, including additional games that were not played in the BO condition due to time constraints, the average benefit for revealers was 64 when agreement was reached, 24 points higher than the benefit in games where agreement was reached without revelation ($SE = 4.3$, $t(146) = -2.32$, $p = 0.01$). As can be expected, this trend was reversed for goal solicitors, whose benefit in the case where goals were revealed was 22, about 30 points lower than in the case when they were not revealed ($SE = 5$, $t(146) = -2.8$, $p = 0.02$).

5.2 Agreement

This section looks at the effect of goal revelation and task dependency on likelihood of players coming to an agreement.

		IBN	
		Agreement	Failure
PBN	Agreement	11	7
	Failure	16	31

Table 5.3: Pairwise agreement ratio

Table 5.3 presents a pairwise comparison between the number of agreements in both conditions. As can be seen in the table, 16 of the games that resulted in agreement in the IBN condition had failed in the PBN condition. In contrast, only 7 of the games that succeeded in the PBN condition failed to reach agreement in the IBN condition, and this difference was statistically significant ($\chi^2(1, N = 65) = 3.92, p = 0.04$).

	Single-dependent	**Co-dependent**	**Total**
IBN	33 (64%)	14 (100%)	47
PBN	26 (50%)	12 (85%)	38

Table 5.4: Agreement frequencies by dependency

Table 5.4 shows the number of games that resulted in agreement as a function of the dependency relationship between players. As can be shown by the table, in both conditions, co-dependent players were significantly more likely to reach agreement than combinations of independent and dependent players (Fisher's exact test $p < 0.01$). In all, there were 9 more acceptances in the IBN condition than there were in the PBN condition (47 compared to 38), and this difference was statistically significant (paired t-test, $p < 0.05$).

Among all the games played in the IBN condition, 32 out of 39 games in which one goal was revealed resulted in agreement, and 7 out of 10 games in which two goals were revealed resulted in agreement ($\chi^2(2) = 7.5, p = 0.02$).

5.3 Inquiry and Revelation behaviour

Among all 113 IBN-played games, there were 39 games in which one goal was revealed, and 10 games in which two goals were revealed, making for a total of 59 revelations. In all, at least one goal was revealed in 43% of the games. The majority of goal revelations (72%) were performed by dependent players, significantly higher than the number of revelations performed by independent players ($\chi^2(1) = 6.1, p = 0.01$).

At the individual player level in the same games, out of the 26 times independent players were asked to reveal their goals, they did so 61% of the time, whereas 85% of the 49 solicitations of dependent players' goal locations were met with a revelation by the dependent player. Overall, players revealed their goals 77.3% of the 75 times they were solicited for the information.

5.4 Analysis Notes

The analysis of the data reveals that when human negotiators were asked to reveal the locations of their goals in response to offers they made, they were more likely than not to reveal this information. Dependent players, in particular, were more likely than independent players to reveal their goals.

Games which featured interest inquiry and revelation dialogue options were more likely to result in agreements than games in which negotiation was based purely on positions. Specifically among games where interest information exchange dialogue was available, players were significantly more likely to agree in games that featured goal revelation(s) compared to those that did not.

The revelation of goals also had a noticeable effect on performance of the players. Games with goal revelations in them resulted in better social outcomes compared to those without. Playing IBN games was also found to be better for the social benefit than PBN games. Dependent players in games where at least one player revealed their goal experienced a significant increase in individual outcome compared to their performance in the same settings played with position-based negotiation.

Independent players, on the other hand, experience a utility decrease in games which featured goal revelation, although they experience a significantly greater overall benefit in interest-based negotiation games compared to their benefit from the same settings played with position-based negotiation protocols. In interest-based negotiation games where no goals were revealed, they did significantly better than they did in the same settings played with position-based negotiation protocols.

Chapter 6

Discussion

The adapted platform proved, for the most part, to be robust and reliable during the experiments. Enough flexibility was built into the software, though, to ensure that the controller settings could be manipulated and the pooled sequence of games could be reloaded and resumed from any point of failure. This feature came in handy during the experiments, as the platform is built to run as a client-server system, and some glitches due to network issues and other technical difficulties are to be expected. This flexibility and robustness will prove to be useful in future experiments.

The experiments themselves were the source of a large amount of data, especially from the IBN games. Being able to identify games with the same original decision-making settings was useful in analysing how the same games played out differently when played with the PBN and IBN protocols. Creating a game pool that could be reused between protocols was also vital to this kind of analysis.

The human subjects appear to show much trust when it comes to goal revelation. While dependent players revealed on asking at a higher rate than dependent players, both classes of players revealed more often than not, when they were asked to do so. Interestingly, in games where goals were known to at least one of the players, independent players actually lost some utility to help the dependent players gain much more utility, raising the social benefit in the process. Therefore, the aforementioned trust may have been justified after all.

However, independent players gained significant utility in IBN games that did **not** feature goal revelation. One possible explanation for this is that, because the dependent players may have volunteered to offer most of or their entire surpluses in each proposal, there was no need on the part of the independent players to ask for goals in those cases. Even though such a situation would be bad for the dependent player, it would be very beneficial for the independent player, almost regardless as to which party know the other's goal.

Chapter 7

Future Work

While much has been revealed about the behaviour and effects on performance of players of goal revelation in mediated negotiation games, the scope of the experimentation was limited to the study of a particular variant of interest-based negotiation protocol. While the game configurations, experimental setups and utility functions used for the experiments were designed to cover as many facets and combinations of gameplay as possible, there are others yet to be explored.

Therefore, a few additional ways in which the experiments could be modified in order to learn more about human behaviour in IBN games are presented below.

- Running an experiment with full initial visibility of opponents' goals, and comparing the data to that which was collected from the IBN experiment.

- Modifying the protocol to allow players to ask about their opponent's goals without having to have been made an offer.

- Modifying the protocol to allow players to reveal their goals without having to be asked to do so first.

- Changing the utility function to remove the distance-to-goal penalty i.e. making it all-or-nothing.

- Modifying the initial player positioning to allow for more than one absolute shortest path to the players' goals.

- Having each human subject play each possible opponent a fixed number of times, with both aware of each other's identities, in order to study active reciprocity.

- Building classifiers based on the collected data. For example, to decide, based on the number of chips one has, and the no-negotiation alternative score one can get, whether it is beneficial to reveal one's goal.

Chapter 8

Conclusions

This work opens avenues of exploration in the study of goal revelation in negotiation. It features the first empirical study of the effects of goal revelation on the individual and social outcomes of negotiations, and of the negotiation behaviour of human experiment subjects. It demonstrates the usefulness of the Colored Trails platform for experimentation on human subjects in negotiation games, and the ability to extract interesting results from such experiments.

The experiments revealed some interesting patterns in human negotiation with the use and consequences of goal revelation for players in various dependency situations. Analysis of the data from the experiments reveals that human negotiators are likely to reveal their goals when the information is solicited, especially if they are dependent on the negotiator requesting the information. For interest-based negotiation games, and especially games featuring goal revelation, the chances of agreement are found to be higher. The revelation of goals also leads to a significant increase in benefit for players in a dependent position, and an increase in the social benefit as well.

The protocols and the experiments developed for this dissertation set a precedent for further study of goal revelation in negotiation, including studies that may employ other platforms. They also lay the groundwork for the development of computer agents that can competently negotiate with humans in mediated settings that feature goal revelation options, empowered with rules derived from and trained on the data collected from experiments like these.

Bibliography

Archer, A., Papadimitriou, C., Talwar, K. and Tardos, E. (2003), An approximate truthful mechanism for combinatorial auctions with single parameter agents, *in* 'SODA '03: Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms', Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, pp. 205–214.

Atran, S., Axelrod, R. and Davis, R. (2007), Sacred barriers to conflict resolution, *in* 'Science', Vol. 317, pp. 1039–1040.

Babaioff, M. and Walsh, W. E. (2003), Incentive-compatible, budget-balanced, yet highly efficient auctions for supply chain formation, *in* 'EC '03: Proceedings of the 4th ACM conference on Electronic commerce', ACM, New York, NY, USA, pp. 64–75.

Bosse, T. and Jonker, C. M. (2005), Human vs. computer behaviour in multi-issue negotiation, *in* 'RRS '05: Proceedings of the Rational, Robust, and Secure Negotiation Mechanisms in Multi-Agent Systems (RRS'05) on Multi-Agent Systems', IEEE Computer Society, Washington, DC, USA, p. 11.

Bosse, T., Jonker, C. M. and Treur, J. (2004), Experiments in human multi-issue negotiation: Analysis and support, *in* 'AAMAS '04: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems', IEEE Computer Society, Washington, DC, USA, pp. 671–678.

Buffett, S. and Spencer, B. (2005), Learning opponents' preferences in multi-object automated negotiation, *in* 'ICEC '05: Proceedings of the 7th international conference on Electronic commerce', ACM, New York, NY, USA, pp. 300–305.

Coehoorn, R. M. and Jennings, N. R. (2004), Learning on opponent's preferences to make effective multi-issue negotiation trade-offs, *in* 'ICEC '04: Proceedings of the 6th international conference on Electronic commerce', ACM, New York, NY, USA, pp. 59–68.

Cramton, P., Shoham, Y. and Steinberg, R. (2004), Combinatorial auctions, Papers of Peter Cramton 04mit, University of Maryland, Department of Economics - Peter Cramton.
URL: *http://ideas.repec.org/p/pcc/pccumd/04mit.html*

Faratin, P., Sierra, C. and Jennings, N. R. (2002), "Using similarity criteria to make issue trade-offs in automated negotiations", *Artificial Intelligence* , Vol. 142, pp. 205–237.

Gal, Y., Dsouza, S., Pasquier, P., Rahwan, I. and Abdallah, S. (2009), The effects of goal revelation on computer-mediated negotiation, *in* 'Proceedings of Cogsci 2009: 31st Annual Conference of the Cognitive Science Society (To be published)', pp. –.

Gal, Y., Grosz, B. J., Pfeffer, A., Shieber, S. M. and Allain, A. (2007), The influence of task contexts on the decision-making of humans and computers., *in* B. N. Kokinov, D. C. Richardson, T. Roth-Berghofer and L. Vieu, eds, 'CONTEXT', Vol. 4635 of *Lecture Notes in Computer Science*, Springer, pp. 206–219.
URL: *http://dblp.uni-trier.de/db/conf/context/context2007.html% GalGPSA07*

Goh, K.-Y., Teo, H.-H., Wu, H. and Wei, K.-K. (2000), Computer-supported negotiations: an experimental study of bargaining in electronic commerce, *in* 'ICIS '00: Proceedings of the

twenty first international conference on Information systems', Association for Information Systems, Atlanta, GA, USA, pp. 104–116.

Grosz, B. J., Kraus, S., Talman, S., Stossel, B. and Havlin, M. (2004), The influence of social dependencies on decision-making: Initial investigations with a new game, *in* 'AAMAS '04: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems', IEEE Computer Society, Washington, DC, USA, pp. 782–789.

Heiskanen, P., Ehtamo, H. and Hamalainen, R. (2001), "Constraint proposal method for computing pareto solutions in multi-party negotiations", *European Journal of Operational Research* , Vol. 133, pp. 44–61.

Hindriks, K. and Tykhonov, D. (2008), Opponent modelling in automated multi-issue negotiation using bayesian learning, *in* 'AAMAS '08: Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems', International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, pp. 331–338.

Hindriks, K. V. and Jonker, C. M. (2008), Creating human-machine synergy in negotiation support systems: Towards the pocket negotiator, *in* W.-P. Brinkman and K. V. Hindriks, eds, 'HuCom 2008: Proceedings of the First International Working Conference on Human Factors and Computational Models in Negotiation', TU Delft, Delft, The Netherlands, pp. 47–54.

Hofstede, G. J., Jonker, C. M. and Verwaart, T. (2008), An agent model for the influence of culture on bargaining, *in* 'HuCom 2008: Proceedings of the First International Working Conference on Human Factors and Computational Models in Negotiation', TU Delft, Delft, The Netherlands, pp. 39–46.

Kakas, A. and Moraitis, P. (2006), Adaptive agent negotiation via argumentation, *in* 'AAMAS '06: Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems', ACM, New York, NY, USA, pp. 384–391.

Katz, R. and Kraus, S. (2007), Gender-sensitive automated negotiators, *in* 'AAAI', pp. 821–826.
URL: *http://www.informatik.uni-trier.de/ ley/db/conf/aaai/aaai2007.html#KatzK07*

Kersten, G. E. (2001), "Modeling distributive and integrative negotiations. review and revised characterization", *Group Decision and Negotiation* , Vol. 10.

Koenig, S., Tovey, C. A., Lagoudakis, M. G., Markakis, E., Kempe, D., Keskinocak, P., Kleywegt, A. J., Meyerson, A. and Jain, S. (2006), The power of sequential single-item auctions for agent coordination., *in* 'AAAI', AAAI Press.
URL: *http://dblp.uni-trier.de/db/conf/aaai/aaai2006.html#KoenigTLMKKKMJ06*

Kraus, S., Sycara, K. and Evenchik, A. (1998), "Reaching agreements through argumentation: A logical model and implementation", *Artificial Intelligence* , Vol. 104, pp. 1–69.

Lai, G., Sycara, K. and Li, C. (2007), A pareto optimal model for automated multi-attribute negotiations, *in* 'AAMAS '07: Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems', ACM, New York, NY, USA, pp. 1–3.

Lin, R., Kraus, S., Wilkenfeld, J. and Barry, J. (2008), "Negotiating with bounded rational agents in environments with incomplete information using an automated agent", *Artif. Intell.* , Vol. 172, Elsevier Science Publishers Ltd., Essex, UK, pp. 823–851.

Loewenstein, G. F., Thompson, L. and Bazerman, M. H. (1989), Social utility and decision making in interpersonal contexts, *in* 'Journal of Personality and Social Psychology', Vol. 57, pp. 426–441.

Loewenstein, J. and Brett, J. (2007), Goal framing predicts strategy revision: When and why negotiators reach integrative agreements, *in* 'Proceedings of the cognitive science society', pp. 443–448.

Narayanan, V. and Jennings, N. R. (2005), An adaptive bilateral negotiation model for e-commerce settings, *in* 'CEC '05: Proceedings of the Seventh IEEE International Conference on E-Commerce Technology', IEEE Computer Society, Washington, DC, USA, pp. 34–41.

Pasquier, P., Hollands, R., Dignum, F., Rahwan, I. and Sonenberg, L. (2007), An empirical study of interest-based negotiation, *in* 'ICEC '07: Proceedings of the ninth international conference on Electronic commerce', ACM, New York, NY, USA, pp. 339–348.

Rahwan, I., Pasquier, P., Sonenberg, L. and Dignum, F. (2007), On the benefits of exploiting underlying goals in argument-based negotiation, *in* 'AAAI', pp. 116–121.
URL: *http://www.informatik.uni-trier.de/ ley/db/conf/aaai/aaai2007.html#RahwanPSD07*

Rahwan, I., Sonenberg, L. and Dignum, F. (2003), Towards interest-based negotiation, *in* 'AAMAS '03: Proceedings of the second international joint conference on Autonomous agents and multiagent systems', ACM, New York, NY, USA, pp. 773–780.

Rubinstein, A. (1985), "A bargaining model with incomplete information about preferences", *Econometrica* , Vol. 53, pp. 1151–1172.

Simon, H. (1957), A behavioural model of rational choice, *in* H. Simon, ed., 'Models of man: social and rational; mathematical essays on rational human behavior in a social setting', J. Wiley, New York, pp. 241–260.

Soo, V.-W. and Hung, C.-A. (2002), On-line incremental learning in bilateral multi-issue negotiation, *in* 'AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems', ACM, New York, NY, USA, pp. 314–315.

Talman, S., Hadad, M., Gal, Y. and Kraus, S. (2005), Adapting to agents' personalities in negotiation, *in* 'AAMAS '05: Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems', ACM, New York, NY, USA, pp. 383–389.

Vorauer, J. D. and Claude, S.-D. (n.d.), "Perceived versus actual transparency of goals in negotiation", *Personality and Social Psychology Bulletin* , Vol. 24, Sage Publications, 2455 Teller Road, Thousand Oaks, CA, USA, pp. 371–385.

Weingart, L. R., Hyder, E. B. and J., P. M. (1996), "Knowledge matters : The effect of tactical descriptions on negotiation behavior and outcome", *Journal of personality and social psychology* , Vol. 70, American Psychological Association, Washington, DC, USA, pp. 1205–1217.

Winoto, P., McCalla, G. and Vassileva, J. (2004), Non-monotonic-offers bargaining protocol, *in* 'AAMAS '04: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems', IEEE Computer Society, Washington, DC, USA, pp. 1072–1079.

Zhang, X., Lesser, V. and Wagner, T. (2002), Integrative negotiation in complex organizational agent systems, *in* 'AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems', ACM, New York, NY, USA, pp. 503–504.

Zlotkin, G. and Rosenschein, J. S. (1996), "Mechanisms for automated negotiation in state oriented domains", *Journal of Artificial Intelligence Research* , Vol. 5, pp. 163–238.