2011

Spatio-temporal Negotiation Protocols

Yi Luo
University of Central Florida

Spatio-temporal negotiation protocols

by

YI LUO

B.S. Huazhong University of Science and Technology, China, 2003
M.S. University of Central Florida, 2009

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the School of Electrical Engineering and Computer Science
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Spring Term
2011

Major Professor: Ladislau Bölöni

ABSTRACT

Canonical problems are simplified representations of a class of real world problems. They allow researchers to compare algorithms in a standard setting which captures the most important challenges of the real world problems being modeled. In this dissertation, we focus on negotiating a collaboration in space and time, a problem with many important real world applications. Although technically a multi-issue negotiation, we show that the problem can not be represented in a satisfactory manner by previous models. We propose the "Children in the Rectangular Forest" (CRF) model as a possible canonical problem for negotiating spatio-temporal collaboration.

In the CRF problem, two embodied agents are negotiating the synchronization of their movement for a portion of the path from their respective sources to destinations. The negotiation setting is zero initial knowledge and it happens in physical time. As equilibrium strategies are not practically possible, we are interested in strategies with bounded rationality, which achieve good performance in a wide range of practical negotiation scenarios. We design a number of negotiation protocols to allow agents to exchange their offers. The simple negotiation protocol can be enhanced by schemes in which the agents add additional information of the negotiation flow to aid the negotiation partner in offer formation. Naturally, the performance of a strategy is dependent on the strategy of the opponent and the

characteristics of the scenario. Thus we develop a set of metrics for the negotiation scenario which formalizes our intuition of collaborative scenarios (where the agents' interests are closely aligned) versus competitive scenarios (where the gain of the utility for one agent is paid off with a loss of utility for the other agent).

Finally, we further investigate the sophisticated strategies which allow agents to learn the opponents while negotiating. We find strategies can be augmented by collaborativeness analysis: the approximate collaborativeness metric can be used to cut short the negotiation. Then, we discover an approach to model the opponent through Bayesian learning. We assume the agents do not disclose their information voluntarily: the learning needs to rely on the study of the offers exchanged during normal negotiation. At last, we explore a setting where the agents are able to perform physical action (movement) while the negotiation is ongoing. We formalize a method to represent and update the beliefs about the valuation function, the current state of negotiation and strategy of the opponent agent using a particle filter.

By exploring a number of different negotiation protocols and several peer-to-peer negotiation based strategies, we claim that the CRF problem captures the main challenges of the real world problems while allows us to simplify away some of the computationally demanding but semantically marginal features of real world problems.

Dedicated to,

My parents

and

My academic advisor

Dr. Ladislau Bölöni

ACKNOWLEDGMENTS

I would like to express my sincere appreciation and gratitude to my advisor, Dr. Ladislau Bölöni for bringing me to the beautiful city of Orlando and providing me an excellent atmosphere to do research. His excellent guidance, caring, patience and infectious enthusiasm to research have been my major driving force throughout the graduate career at the University of Central Florida. I appreciate Dr. Gary T. Leavens, Dr. Damla Turgut and Dr. Yue Zhao for serving as members of my dissertation committee. Their suggestions and advices have helped me immensely to prepare the dissertation. I thank my parents for their unflagging support to my oversea study. Last but not least, I thank my friends, my roommates and my labmates. Their company presents me a cherished memory for my life.

This dissertation was partially funded by NSF Information and Intelligent Systems division under award 0712869 and the Army Research Laboratory under Cooperative Agreement Number W911NF-06-2-0041.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

The main topic of this dissertation is collaboration in space and time between embodied agents, a subject with many real world applications. Finding an optimal solution from the point of view of an outside observer is computationally difficult and for many applications is unrealistic due to the conflicting interests of the agents. Thus, instead of a centralized solution we rely on agents negotiating an agreement to collaborate. Such a solution will not be optimal, but the self-interested nature of the agents will motivate them to find the solution which is satisfactory by both of them. In this dissertation we propose a canonical problem to study spatio-temporal negotiation. We develop a series of negotiation protocols, negotiation strategies and techniques to evaluate the scenarios and the performance.

The remainder of this chapter is organized as follows. We introduce spatio-temporal collaboration and provide real world applications in Section 1.1. We motivate the use of negotiation for establishing spatio-temporal collaboration in Section 1.2. We analyze the characteristics of spatio-temporal negotiation in Section 1.3. At last, we discuss the contributions and organization of the dissertation in Section 1.4 and Section 1.5.

1.1 Spatio-temporal collaboration

Collaboration between embodied agents often requires spatial and temporal collocation. Agents need to coordinate their movements, agree on meeting locations, time, common path and speed, as well as locations where they split and start moving on independent trajectories. Such problems appear as sub-problems in many practical applications such as the transportation system and teamwork applications. In the following, we provide some real world examples which require spatio-temporal collaboration.

1.1.1 Planning schedules in the transportation system

Planning the schedules is a challenging task for today's transportation companies. The schedules need to optimize the profit of the company while respecting regulations, requirements and resource limitations.

An example is bus transportation lanes, either local (such as LYNX in Orlando) or inter-city transportation lanes such as Greyhound (see Figure 1.1). To plan a schedule, the administrators need to decide how many shuttles they will send from one location to the others and how many intermediate stations these shuttles will stop along the paths. In addition, they should also decide the departure times and the approximate arrival time at the intermediate stations. Figure 1.2 shows one of the schedules in 1949 from Chicago to Louisville. The schedules must be adapted according to the flow of passengers. For

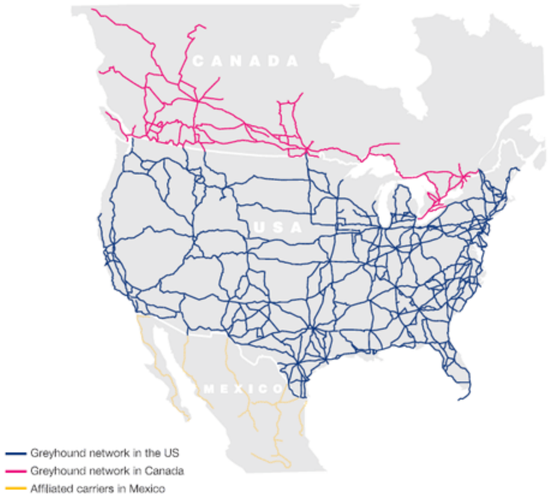
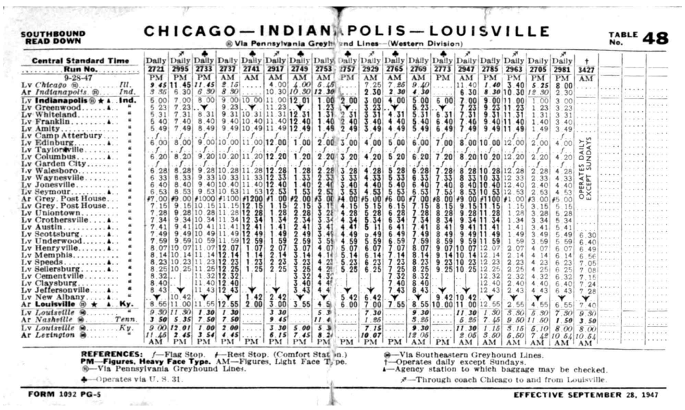


Figure 1.1: The transportation network for Greyhound Lines.



SOUTHBOUND READ DOWN

CHICAGO—INDIANAPOLIS—LOUISVILLE

Ⓡ Via Pennsylvania Greyhound Lines—(Western Division)

TABLE No. 48

Central Standard Time	↗	♣		♣	↗	♣		♣	↗	♣		♣	↗		↗		↗	†
	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Daily	
Run No.	2721	2995	2733	2737	2741	2917	2749	2753	2757	2929	2765	2769	2773	2947	2785	2963	2705	2981	3427
9-28-47	PM	PM	PM	AM	AM	AM	AM	AM	PM	AM	AM	AM	PM	AM	PM	PM	PM	PM	AM
Lv Chicago Ⓡ … Ill.	9.45	11.45	11.45	2.15	....	4.00	4.00	5.45	....	7.25	7.25	9.40	....	11.40	1.40	3.40	5.25	8.00	
Ar Indianapolis Ⓡ … Ind.	3.35	6.30	6.30	8.30	....	10.30	10.30	12.30	....	2.30	2.30	4.30	....	6.30	8.30	10.30	12.30	2.30	
Lv Indianapolis Ⓡ ★ ▲ Ind.	5.00	7.00	8.00	9.00	10.00	11.00	12.01	1.00	2.00	3.00	4.00	5.00	6.00	7.00	9.00	11.00	1.00	3.00	
Lv Greenwood ▲ "	5.23	7.23	↓	9.23	↓	11.23	↓	1.23	↓	3.23	↓	5.23	↓	7.23	9.23	11.23	1.23	3.23	
Lv Whiteland "	5.31	7.31	8.31	9.31	10.31	11.31	12.31	1.31	2.31	3.31	4.31	5.31	6.31	7.31	9.31	11.31	1.31	3.31	
Lv Franklin ▲ "	5.40	7.40	8.40	9.40	10.40	11.40	12.40	1.40	2.40	3.40	4.40	5.40	6.40	7.40	9.40	11.40	1.40	3.40	
Lv Amity "	5.49	7.49	8.49	9.49	10.49	11.49	12.49	1.49	2.49	3.49	4.49	5.49	6.49	7.49	9.49	11.49	1.49	3.49	
Lv Camp Atterbury "	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	
Lv Edinburg ▲ "	6.00	8.00	9.00	10.00	11.00	12.00	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	10.00	12.00	2.00	4.00	OPERATES DAILY EXCEPT SUNDAYS
Lv Taylorsville "	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	
Lv Columbus ▲ "	6.20	8.20	9.20	10.20	11.20	12.20	1.20	2.20	3.20	4.20	5.20	6.20	7.20	8.20	10.20	12.20	2.20	4.20	
Lv Garden City "	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	
Lv Walesboro "	6.28	8.28	9.28	10.28	11.28	12.28	1.28	2.28	3.28	4.28	5.28	6.28	7.28	8.28	10.28	12.28	2.28	4.28	
Lv Waynesville "	6.33	8.33	9.33	10.33	11.33	12.33	1.33	2.33	3.33	4.33	5.33	6.33	7.33	8.33	10.33	12.33	2.33	4.33	
Lv Jonesville "	6.40	8.40	9.40	10.40	11.40	12.40	1.40	2.40	3.40	4.40	5.40	6.40	7.40	8.40	10.40	12.40	2.40	4.40	
Lv Seymour ▲ "	6.53	8.53	9.53	10.53	11.53	12.53	1.53	2.53	3.53	4.53	5.53	6.53	7.53	8.53	10.53	12.53	2.53	4.53	
Ar Grey. Post House "	ƒ7.00	ƒ9.00	ƒ1000	ƒ1100	ƒ1200	ƒ1.00	ƒ2.00	ƒ3.00	ƒ4.00	ƒ5.00	ƒ6.00	ƒ7.00	ƒ8.00	ƒ9.00	ƒ1100	ƒ1.00	ƒ3.00	ƒ5.00	
Lv Grey. Post House "	7.15	9.15	10.15	11.15	12.15	1.15	2.15	3.15	4.15	5.15	6.15	7.15	8.15	9.15	11.15	1.15	3.15	5.15	
Lv Uniontown "	7.28	9.28	10.28	11.28	12.28	1.28	2.28	3.28	4.28	5.28	6.28	7.28	8.28	9.28	11.28	1.28	3.28	5.28	
Lv Crothersville ▲ "	7.34	9.34	10.34	11.34	12.34	1.34	2.34	3.34	4.34	5.34	6.34	7.34	8.34	9.34	11.34	1.34	3.34	5.34	
Lv Austin ▲ "	7.41	9.41	10.41	11.41	12.41	1.41	2.41	3.41	4.41	5.41	6.41	7.41	8.41	9.41	11.41	1.41	3.41	5.41	
Lv Scottsburg ▲ "	7.49	9.49	10.49	11.49	12.49	1.49	2.49	3.49	4.49	5.49	6.49	7.49	8.49	9.49	11.49	1.49	3.49	5.49	6.30
Lv Underwood ▲ "	7.59	9.59	10.59	11.59	12.59	1.59	2.59	3.59	4.59	5.59	6.59	7.59	8.59	9.59	11.59	1.59	3.59	5.59	6.40
Lv Henryville ▲ "	8.07	10.07	11.07	12.07	1.07	2.07	3.07	4.07	5.07	6.07	7.07	8.07	9.07	10.07	12.07	2.07	4.07	6.07	6.49
Lv Memphis "	8.14	10.14	11.14	12.14	1.14	2.14	3.14	4.14	5.14	6.14	7.14	8.14	9.14	10.14	12.14	2.14	4.14	6.14	6.56
Lv Speeds ▲ "	8.23	10.23	11.23	12.23	1.23	2.23	3.23	4.23	5.23	6.23	7.23	8.23	9.23	10.23	12.23	2.23	4.23	6.23	7.05
Lv Sellersburg ▲ "	8.25	10.25	11.25	12.25	1.25	2.25	3.25	4.25	5.25	6.25	7.25	8.25	9.25	10.25	12.25	2.25	4.25	6.25	7.08
Lv Cementville "	8.32	↓	11.32	12.32	↓	↓	3.32	4.32	↓	↓	7.32	8.32	↓	↓	12.32	2.32	4.32	6.32	7.15
Lv Claysburg "	8.40		11.40	12.40			3.40	4.40			7.40	8.40			12.40	2.40	4.40	6.40	7.24
Lv Jeffersonville ▲ "	8.43	↓	11.43	12.43	↓	↓	3.43	4.43	↓	↓	7.43	8.43	↓	↓	12.43	2.43	4.43	6.43	7.28
Lv New Albany ▲ "	↓	10.42	↓	↓	1.42	2.42	↓	↓	5.42	6.42	↓	↓	9.42	10.42	↓	↓	↓	↓	↓
Ar Louisville Ⓡ ★ ▲ Ky.	8.55	11.00	11.55	12.55	2.00	3.00	3.55	4.55	6.00	7.00	7.55	8.55	10.00	11.00	12.55	2.55	4.55	6.55	7.40
Lv Louisville Ⓢ "	9.30	11.30	1.30	1.30	....	3.30		5.30	....	7.30	....	9.30	....	11.30	1.30	3.30	5.30	7.30	9.30
Ar Nashville Ⓢ … Tenn.	3.50	5.35	7.50	7.50	....	9.45		11.4[illegible]	....	1.25	....	3.25	....	5.25	7.45	9.50	11.50	1.50	3.50
Lv Louisville Ⓢ … Ky.	9.00	12.01	1.00	2.00	....	3.30	5.00	5.3[illegible]	....	7.15	....	9.30	....	11.30	1.15	3.15	5.10	8.00	8.00
Ar Lexington Ⓢ "	11.45	2.45	3.54	4.45	....	6.15	7.45	8.2[illegible]	....	10.07	....	12.05	....	2.05	3.50	5.50	7.42	10.54	10.54
	AM	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	AM	PM	AM	AM	AM	AM	AM	AM

REFERENCES: f—Flag Stop. ƒ—Rest Stop. (Comfort Station.) PM—Figures, Heavy Face Type. AM—Figures, Light Face Type. Ⓡ—Via Pennsylvania Greyhound Lines. ♣—Operates via U. S. 31.

Ⓢ—Via Southeastern Greyhound Lines. †—Operates daily except Sundays. ▲—Agency station to which baggage may be checked. ↗—Through coach Chicago to and from Louisville.

FORM 1092 PG-5

EFFECTIVE SEPTEMBER 28, 1947

Figure 1.2: The Greyhound schedule from Chicago to Louisville on September 28, 1947.

example, during the winter, people like to travel south to Florida. In the summer, people prefer to travel late in the day. The schedulers should also concern about the transfer of passengers to connecting shuttles, which requires the synchronization between shuttles, effectively establishing meeting points.

Similar problems are encountered by mail delivery companies such as UPS and FedEx, which need to dynamically adjust their schedules according to the demand which is variable and highly seasonal.

As long as this class of problems appears in the context of a single company, where there is no conflict of interest between the participating shuttles and delivery trucks, the problem is solved using optimization and operations research techniques, rather than negotiation. There are, however, many instances where the transportation problems are solved in the context of multiple, self-interested entities. For example, transportation companies might use the services of local contractors. Transfer of passengers between regional transport lines must be facilitated. These challenges can only be solved by resolving the conflicting interests of the participants, that is, through negotiation.

1.1.2 Convoy formation by vehicles

Convoy formation is a widely deployed method of collaboration between embodied agents such as motor vehicles or ships. The members of the convoy are traveling together for mutual

Figure 1.3: US Navy warships escort the tanker Gas King in 1987

support and protection. Often, a convoy is organized with armed defensive support, though it may also be used in a non-military sense, for example when driving through remote areas.

Naval convoys (see Figure 1.3), for example, have been used for hundreds of years, and examples of merchant ships traveling under naval protection have been traced back to the 12th Century. According to literatures, by the French Revolutionary Wars of the late 18th century, effective naval convoy tactics had been developed to ward off pirates and privateers.

To form a convoy, the participants should keep themselves in the communication range. They need to coordinate their movements, traveling to rendezvous under the time constraints. During the convoy, they need to agree on a common path and speed, and they also need to find locations where they split such that they can continue their independent trajectories.

1.2 Negotiating collaboration in space and time

The examples in the previous section were, at least theoretically, solvable using centralized solutions, as the interests of the participants were largely common. There are many other examples where the agents need to collaborate in space and time but the participants are self-interested and have private information which they don't want to share. In these cases a global coordinator or a pre-planned solution is not applicable. The collaboration can be agreed upon through negotiation, with the agents exchanging offers and agreeing upon a solution which is acceptable to all of them. In the following, we provide two examples in the spatio-temporal collaborations which favor negotiation.

1.2.1 Transportation for elderly and disabled persons

Many local transportation companies in the United States are providing door-to-door transportation services for the elderly or disabled persons who can not use the fixed route bus service. For instance, in Orlando, the ACCESS LYNX program is providing more than 3100 scheduled passenger trips per day, using a large number of shuttle type vehicles. The vehicles might be operated by external contractors. Naturally, these services can not follow the one-person/one-trip model followed by taxis, as that would be prohibitively expensive. Requests for transportation are submitted by phone by the passengers. These requests need to be satisfied using the shuttles currently in service. The shuttles need to organize their path and schedule dynamically, such that they provide the best possible service. An incoming request modifies the path of the shuttle, which needs to make a detour to pick up the passenger. The transfer of passengers from one shuttle to another needs to be scheduled dynamically, the rendezvous of the shuttles of the transfer point agreed upon.

Let us now envision a negotiation-based solution to the problem of efficient scheduling of the passenger transportation. This assumes that the shuttles are competing with each other for business, using performance measures such as total passenger miles, total number of passengers served or total passenger miles / total miles. In addition, the goals of the dispatcher use different, global performance measures, such as average time before pickup or average time to destination.

One way to organize the negotiation process is to allow only pairwise negotiations between the dispatcher and the shuttle. To satisfy a new transportation request, the dispatcher might contact several shuttles, and negotiate a modification of the route in order to pick up a new passenger. If the transportation request can not be satisfied with a single shuttle, the dispatcher might negotiate a rendezvous of two shuttles in order to arrange for a transfer of the passenger. This is an example of a co-negotiation; the offers of the dispatcher in one negotiation are conditioned by the evolution of the other negotiation process. Finally, even previous agreements can be revisited based on the set of new requirements.

The issues under negotiation can be described with spatio-temporal constraints; for instance, an offer might look like this: "pick up a passenger at location L_1 after time t_1, drop him/her at location L_2 before time t_2 but do not leave location L_2 before time t_3".

More complex negotiation patterns can also be deployed. For instance, the drivers might be able to negotiate directly among themselves, the passengers might get involved in the negotiation as well, and the negotiation might include incentives and dis-incentives as well.

1.2.2 Convoy formation in disaster response applications

Efficient response in face of natural disasters, such as Hurricane Katrina in New Orleans, the Asian tsunami or the earthquake in Pakistan, requires participants to form teams and coordinate their actions. In the immediate aftermath of a disaster previously safe areas might turn into unsafe or inaccessible. The environment might contain new sources of danger in the form of natural obstacles (damaged buildings) or even hostile agents (such as looters or stray dogs).

The tasks facing the rescue teams appear unpredictably. The discovery of a wounded person at a dangerous location creates a new task with specific logistics, protection and medical facets.

The organization of the rescue teams can not be pre-planned, and more often than not, a centralized coordination is not possible. For instance, in case of Hurricane Katrina, the central dispatcher unit of the police was flooded; the police could use their radios only as short-distance walkie-talkies. Furthermore, although some of the disaster management teams are pre-established, trained together and have a clear pattern of command and control, many teams are assembled on an ad hoc basis, as a response to emerging tasks. Teams are composed from heterogeneous groups of entities: persons, vehicles, service animals, and so on. Team members might not report to the same chain of command, might have communication problems and their interests might not be completely aligned. For instance, the state police and guerilla groups might cooperate in a rescue operation but resume hostilities after the emergency.

Thus, the organization of disaster response activities requires negotiation between agents with different interests. Khan and Bölöni [KB06, BKT06] explored the topic of negotiating convoy formation in disaster response applications. The assumption for this problem is that agents have tasks associated with geographic locations, and in order to achieve those, they need to traverse areas which are accessible only to convoys, but not to individual agents. The negotiation between agents is concerned about temporal commitments regarding specific locations. For instance, in order to successfully join a convoy at location L_{join} the agent will make a commitment to reach that location before time t_{join}, while the convoy will make the commitment that it will leave that location only *after* time t_{join}. To allow the agent to plan ahead towards its task, the convoy takes the commitment that it will reach the preagreed location L_{split} before t_{split}. As the convoy will carry a set of commitments towards all its members, these commitments need to be taken into consideration when new agents are joining the convoy. Naturally, not every commitment is feasible, and the feasibility of a set of commitments needs to be evaluated together.

These problems are only two examples from the much wider class of problems which involve negotiation about collaborative actions in space and time. For instance, the act of passing in soccer (human or robotic) requires the players to agree on the trajectory of the ball, and the future location of the receiver player at a specific time. The act of carrying a piano on the stairs requires the carriers to agree on specific forces to be applied at specific locations and moments.

We claim that negotiation about collaborative actions in space and time is a large class of problems with important practical applications. In the next section, we argue that these negotiation problems can not be adequately modeled by the split the pie game, a worth-oriented negotiation problem which has been widely studied.

1.3 Defining characteristics of spatio-temporal negotiation

One of the canonical problems for agent negotiation is the “split the pie” game [BOR92, OR94] where the participants are negotiating over the partitioning of a pie. The game can be extended in a straightforward game to cover more complex issues. Multi-issue negotiations can be handled by having to split multiple pies, the agents total utility being a function of the pie shares. For reasons related to the computational complexity, the utility function is commonly represented by a weighted sum over the pie shares received by each agent. The agents might or might not know the utility function of their negotiation partner, thus various complete and incomplete information scenarios can be represented. Negotiations with deadlines are represented by imposing a limit on the negotiation rounds. Another, frequently considered aspect is the discount factors, the cost of extended negotiation is represented by the pies shrinking after every negotiation round with a factor of δ [Rub82].

Let us investigate the main reasons why the split the pie game can not serve as a valid canonical problem for negotiations concerning spatio-temporal collaborations. Although, there are many immediate differences in the formulations of the problems, not all of these are fundamental. For instance, our problem domain involves collaboration, while the split the pie model apparently involves a radical conflict of interests. This difference however, is only superficial. While the single issue "split the pie" is a zero-sum game, the multiple pie games are not, because different agents can have different valuations of the different pies, and thus they can reach deals which are advantageous to both of them. With an appropriate utility function, the split the multiple pies game can be used to model the negotiation of collaborative activities. There are however, more fundamental differences, from which we highlight the following five:

(1) Heterogeneous types of issues.

(2) Non-monotonic valuation of issues.

(3) Evolving world (vs. discount factors)

(4) Offers need to be verified for feasibility

(5) Interaction between the negotiation time and physical time

Let us discuss these characteristics in more detail, by contrasting them to the set of negotiation problems modeled by the split the pie game.

(1) Heterogeneous types of issues and **(2) Non-monotonic valuation of the issues**

For the multi-issue split the pie game, all the issues are represented by a numerical value in the $[0, 1]$ interval. There is an assumption that all the different pies have an intrinsic, positive value; the ultimate goal of the negotiation partners being to acquire 100% of all the pies. Of course, the different agents might have different valuations for the different pies, and in a stretch, the utility function might be a non-linear function of the shares[1]. The issues in a split the pie game can be therefore characterized as *worth* values. It makes perfect sense to define the partial derivative of the utility function of agent a with respect to every component of the offer vector. All these partial derivatives will be non-negative, as the game assumes that the utility of a pie can not be negative.

$$\frac{\partial U^a([x_1 \ldots x_n])}{\partial x_i} \geq 0 \quad \forall i \in \{1 \ldots n\} \tag{1.1}$$

That is, the utility function of the agent in the split the pie game is monotonic in all the components of an offer. In fact, when the utility is a linear combination, the partial derivative will be a constant, and exactly the corresponding weight in the linear combination of utilities:

[1] Although most research studies consider the utility to be an additive, linear combination of the values.

$$\frac{\partial U^a([x_1 \ldots x_n])}{\partial x_i} = K_i^a \tag{1.2}$$

However, the situation is different for the case of negotiating spatio-temporal collaboration. Here, the issues under negotiation (that is, the components of an offer) can represent either (a) worth, (b) time values, and (c) points in the 2-dimensional or 3-dimensional space. For the worth-type values, the monotonicity considerations still apply. Things are somewhat more complicated for time values. If the time value represents, for instance, the arrival time to a destination, and we state that it is the goal of the agent to arrive as early as possible, the time value can be immediately mapped into a worth-type issue. However, if the time represents the time of a rendezvous (for instance, catching an airplane), the contribution of the issue to the utility corresponds to a step function: any value smaller than the target has the same value, while every value later than the target is worth 0.

The situation is even more complex for the spatial values. Although there are instances in which a location can be mapped to a worth value (for instance, considering the distance to the final destination), this worth value can not represent the point in the negotiation. Two agents can not agree to rendezvous at "200 miles from New-York", they need to decide on a specific location. There is no objective, positive or negative value in a certain rendezvous point, its value becomes evident only in the context of the remainder of the offer (the time of the rendezvous, the path of the convoy after the rendezvous and so on).

(3) Evolving world (vs. discount factors)

Many negotiation models consider that the utility of a certain offer depends on the moment in the negotiation process when it was presented. Most studies of the split the pie game consider that the value of the issues under negotiation decreases in time, this feature being modeled with discount factors δ, which shrink the pie after every negotiation round. This is a good model of many (but not all[2]) practical situations, and it has the analytical advantage that it models the incentive of the participants to conclude a negotiation. In the case of negotiating spatio-temporal collaboration, we have to consider that (a) the agents may be moving during the negotiation and (b) the time passes. That is, the value of the offer depends on the current location of the agent, as well as the current time. Note that this can not be modeled with discount factors; the value of the offer does not necessarily decrease in time. For instance, the value of a rendezvous point increases if the agent moves on a path which takes it closer to the proposed point, and it starts to decrease once the agent passes the closest location and the distance increases.

(4) Feasibility of the offers

In the split the pie model every correctly formed offer is feasible. However, for spatio-temporal collaborations, there are offers which, although potentially of high value, are not feasible because of the physical world limitations. For instance, one of the participants might propose a rendezvous at a location and time which can not be made by the other participant. The feasibility of an offer can not be evaluated in advance, as it is dependent on the current state of the world. In incomplete information settings, the feasibility needs to

[2]Even in purely worth oriented domains, it is possible that the value of the "pie" increases during the negotiation, consider for instance negotiations concerning real-estate deals.

be verified by all participants separately. Verifying whether an offer is feasible or not can be computationally expensive, as it might involve path planning and estimation of the future state of the world. This has a significant impact in the offer generation step, as the offerer needs to evaluate and verify for feasibility the offers before making them.

(5) The interaction between the negotiation time and physical time

The shrinking pie model abstracts away the physical time, and replaces it with the discrete time model of the negotiation turns. This is a very powerful feature of the model, and a major help in analysis. However, in the class of problems considered by us, we can not make this simplifying step. As we have shown in point (3) above, the agents are acting in an evolving physical world con-committed with the negotiation process. The time taken for the negotiation, including the overhead of offer exchanging and the computational time to generate and evaluate the offers have a direct impact on the outcome of the negotiation. For instance, in fast, real time applications, such as Robocup soccer there is simply no time for exchanging and evaluating multiple offers. In fact, the real-world soccer is increasingly moving towards pre-trained "set pieces", showing that at the speed of human path planning there is no time for evaluating even a single offer - the only offers which can be made are the ones whose values are pre-calculated through previous agreements in the training sessions!

Even if there is time to evaluate several offers, the outcome of negotiation will be different for agents with slow or fast computational facilities (software and/or hardware), and naturally, the outcome is different if one of the agents has more powerful computational facilities than the other.

We conclude that the class of problems representing negotiation about spatio-temporal collaboration have a series of features which are not correctly represented by the "split the pie" model. We are looking for a canonical problem which reflects these features, and at the same time is as simple as possible.

1.4 Contributions

This dissertation makes the following contributions to the theory and practice of spatio-temporal negotiation.

1 We describe the general convoy formation problem (Section 3.1) and propose the "Children in the Rectangular Forest" (CRF) model (Section 3.2) as a possible canonical problem for negotiating spatio-temporal collaboration. We argue that (a) it represents many of the fundamental aspects of this class of problems and (b) it is simple enough to serve as the foundation of formal analysis (Section 3.3).

2 We investigate several methods to evaluate the difficulty of a scenario and develop metrics to match with the intuitions of collaborative scenarios where the agents' interests are closely aligned, versus competitive scenarios where the gain of the utility for one agent is paid off with a loss of utility for the other agent (Chapter 4).

3 We design series of negotiation strategies for a simple negotiation protocol (Section 5.1 and 6.1), and continue to improve these strategies in an argumentation based

negotiation protocol where the simple protocol is enhanced by schemes in which the agents add additional information of the negotiation flow to aid the negotiation partner in the offer formation (Section 5.2 and 6.2).

4 We improve negotiation strategies with real-time opponent modeling. At first, we augment strategies with collaborativeness analysis where the agent can cut short the negotiation in less collaborative scenarios (Section 7.1). Then, we describe an approach for learning the opponent's model based on exchanged offers during normal negotiation (Section 7.2).

5 We formalize the setting for acting while negotiating where embodied agents are able to perform physical action (movement) while the negotiation is ongoing. Then, We develop a method to represent and update the beliefs about the utility function, the current state and strategy of the the opponent agent using a particle filter (Chapter 8).

1.5 The organization of the dissertation

The dissertation is organized in the following way. In Chapter 2, we present the related work. In Chapter 3, we introduce the proposed problem. In Chapter 4, We develop a set of metrics to evaluate the negotiation scenario. Metrics introduced there can be viewed as the criteria to evaluate negotiation strategies. In chapter 5, we design a number of negotiation protocols to allow agents to exchange offers. In Chapter 6, we introduce some strategies applied by the agent in the negotiation. In Chapter 7, we investigate the strategies which allow agents to learn the opponents while negotiating. In Chapter 8, we discuss a setting where agents are able to move while the negotiation is ongoing. At last, we discuss the future work and conclude in Chapter 9.

CHAPTER 2
RELATED WORK

Multiagent system can be viewed as a branch of Artificial Intelligent (AI) while focusing on the design of agents to exhibit intelligent behavior, like human. In [WJ94], agents are described as intelligent entities who take active roles and perform actions that affect the environment, rather than passively allowing their environment to affect them. Like human, agents have belief, desire and intention. They should act *autonomously* and *rationally*. Autonomy means the agents make decision without direct intervention or guidance. Rationality means the ultimate goal of the agents is to maximize the performance with respect to the valuation function.

In the environment with multiple autonomous agents [JFL01], *negotiation* is the most important approach to cooperate and reach agreement. Agents negotiate based on *negotiation protocols*: a set of rules that govern the interaction; they bargain the *negotiation issues*: the range of objects over which agreement must be reached; and they make decision independently to achieve their objectives.

While automated negotiation generated a lot of interest in recent years, negotiation about spatio-temporal issues in embodied agents has received relatively little attention. Nevertheless, many research results in multi-issue negotiation or collaborative robotics have relevance to our work. In the following, we briefly review some of the recent papers which overlap in some aspects with our work, and whose approach (even when applied in a different domain) influenced our approach.

Sandholm and Vulkan [SV99] analyzed the problem of negotiation with internal deadlines where the deadlines are private information of the agents. The negotiated problem is the "split a single pie", zero-sum negotiation. They found that for rational agents, the sequential equilibrium is a strategy which requires agents to wait until their deadline, and at that moment, the agent with the earliest deadline concedes the whole cake.

Fatima, Wooldridge and Jennings [FWJ06, FWJ02] extensively studied the problem of multi-issue negotiation under deadlines. The considered problem is the "split multiple pie problem" where the pies are assumed to be shrinking at each negotiation round. Under both complete information and incomplete information assumptions, the authors compared three negotiation procedures: the package deal procedure where all the issues are discussed together, the simultaneous procedure where issues are discussed independently but simultaneously, and the sequential procedure where issues are discussed one after another. The authors showed that the package deal is the optimal procedure for both agents.

Most of the work above are concerned with the negotiation in the worth oriented domain, where the most frequently taken approach for modeling the overlap between the negotiation time and physical time is the *deadline model*. In these cases, agents do not need a separate action strategy, but they need to take the time in consideration in their negotiation strategy. In [LB07], we found that spatio-temporal negotiation can not be modeled with a discount factor, as the negotiation time and physical time interact together. In [LB09a], we allowed agents to move while negotiating, thus the value of the offers depends on the current physical location as well as the current physical time.

Ito, Klein and Hattori [IKH08] considered negotiations in real world settings where the utility values are non-linear. For instance, the value of the tires and the value of the engines can not be simply added up when designing a car, as the issues constrain each other. The authors proposed an auction-based multi-issue negotiation protocol for negotiating among agents with a non-linear utility settings. The protocol also includes a mediator, which is responsible to choose the deal with the largest social utility from the deals made possible by the bids of the agents.

Golfarelli et al. [GD97] considered the case of robotic agents which are assigned a set of tasks which are attached to physical locations. The tasks carry precedence constraints (execute one specific task earlier than the other) and object constraints (fetch the object in order to execute the task). Agents need to determine, on a network of places and routes, a sequence of places to be visited in order to carry out a set of tasks. Through swapping tasks based on announcement-bid-award mechanism, the agents can decrease their tasks execution costs in the map. An extended version of their work [GR00], allowed the agents to exchange clusters of tasks to avoid being stuck in local minimal. To cluster similar tasks, the authors calculate spatial distance and temporal distance of tasks, and apply thresholds to differentiate between near and far tasks.

In addition to theoretical studies of the multi-agent system, many methodologies are investigated to improve the performance. Most of them are concerned with analyzing the scenarios, designing negotiation protocols and optimizing negotiation strategies.

Brams [Bra90] suggested the *strategy approach* to solve the problem in the multi-agent system. The author claimed that the key element is to design a concession mechanism that let negotiation converge through a series of offer and counter offers. Young [You75] investigated Zeuthen's bargaining model under the condition of bilateral monopoly. In the model, players calculate the maximum probability of conflict they would be willing to accept in preference to acquiescing on the opponent's current offer. The player who has a lower *willingness to risk conflict* will make the next concession.

Most negotiation strategies are designed based on monotonic concession protocol [End06], that is, agents initially makes a proposal that is particularly beneficial to themselves and then incrementally revise their earlier proposals in order to come to an agreement. However, Rahwan et al. [RRJ03] found they can be improved in terms of the likelihood and quality of an agreement by exchanging arguments between agents. They further pointed out that exchanging argument is sometimes essential when various assumptions about agent rationality cannot be satisfied. The most challenges are argument evaluation, generation and selection.

Kakas and Moraitis [KM06] implemented an argumentation based negotiation protocol in a buying-selling scenario. In their framework, if an agent does not satisfy its own goal, it can consider in a conciliation phase, the other agents goal and search for conditions under which it could accept it. During the negotiation, agents exchange arguments for the previous offer. These supporting information are collected and build gradually, thus allowing a form of incremental deliberation of the agents.

Saha and Sen [SS06] discussed the problem of negotiating efficient outcomes in a multi-issue negotiation where some of the parameters of the agent are not common knowledge. The "distributive" and "integrative" scenarios proposed by them are the equivalents of the "competitive" and "collaborative" scenarios we define for the spatio-temporal negotiation problem in [LB08a, LB09b].

Our another concern is the complete/incomplete information for the valuation and feasibility of the offers. In the worth oriented negotiation, the agent might not know the opponent's utility function, nor the strategy it uses. However all the offers are feasible. It is not true in spatio-temporal negotiations. The agent may propose an offer which the opponent can not achieved. Even worse, the agents are not sure the offer they send to the opponent represent a concession for the opponent point of view, especially when the opponent's utility function is non-linear.

To identify these uncertainty, cooperative learning [PL05] can be applied in the agent's strategies. Garland and Alterman [GA04] proposed a learning technique where agents can behave according to the previous experience. In their model, agents store the past successes (in the form of execution trace segments) into entries of memory and apply them to make decisions in the future. Reinforcement learning [KLM96, Tan93] is another approach when the agents don't know if the previous decision is correct or not. All their simple decisions are rewarded based on the output of negotiation. Thus reinforcement learning is only valid when the agent negotiates with the same opponent for many times.

Tykhonov and Hindriks [Dmy08] used Bayesian learning to study the opponent's preference for different issues. They formed a series of discrete hypotheses about the type of the opponent and associated them with probabilities. These probabilities are updated based on Bayes' rule: the difference between the expected bid and the actual bid by the opponent will be used to calculate the probabilities of opponent type in the next round.

Crawford and Veloso [CV07] applied the "experts" algorithm to solve the multi-agent scheduling problem. In their algorithm the agent is helped by a number of "experts", but it needs to decide which experts' advice it should follow. The learning agent can dynamically change its strategy according to its opponents' behavior. The performance is measured in terms of total utility achieved over each of the trials.

Ficici and Pfeffer [FP08] investigated how to simultaneously model agents' preference and their beliefs about others' preference. The proposed model distinguishes three factors: (a) the agent's own utility function, (b) the agent's belief about another agent's utility function, and (c) the agent's belief about how other agent may interact with yet other agents. Gmytrasiewicz and Doshi [GD05] proposed a similar framework where agents maintain beliefs over the physical environment as well as the other agents who may also learn. They called it Interactive POMDPs and demonstrated it has similar properties with POMDPs (convergence of value iteration, the rate of convergence, and piece-wise linearity and convexity of value function). Since the representation of the physical state and the beliefs of other agents is tremendously large, their extended work [DG05] focused on the state estimation procedure using interactive particle filters.

In [LB08b], we applied Bayesian learning during the negotiation and asked the agent to update the belief about opponent's preference based on the last response. In [LB09a], we combined the opponent's physical state and opponent's strategy model in the learning agent's belief. We applied particle filter to accelerate the learning process.

CHAPTER 3
THE PROPOSED PROBLEM

Canonical problems allow researchers to compare algorithms in a standard setting which captures the most important challenges of the real world problems being modeled. Such examples are the block world for planning problems or two-player games for algorithms which learn the behavior of the opponent agent [CM96]. Canonical problems are close relatives to the standardized test beds used in AI research, and frequently, the implementation of the test bed follows a canonical problem. The test bed based controlled experimentation approach had generated controversies [HPC93], with arguments which are just as well applicable to the more theoretical canonical problems as well. Ultimately, the main danger is that the researchers are focusing on problems which are particular only to the testbed, with little relevance to the real world. While a valid criticism, this observation should only make us more careful on selecting our canonical problems, such that they are indeed representative of the real world challenges they represent. Several current initiatives such as the trading agent competition or the Robocup Rescue Simulation League are positioning themselves towards a more accurate modeling of real world problems.

The features which make the split the pie game a good canonical problem is that it is representative of a large class of real world applications. By its simplifying assumptions, it enables a formal analysis of the different components of the negotiation process: the negotiation procedure, the negotiation protocol, the strategies deployed by the negotiation partners, their preferences over the outcomes—usually represented by their utility function and so on. Furthermore, through reducing negotiation problems to the split the pie model the fundamental identity of some negotiation problems can be revealed (which might not be immediately obvious in their original formulation). In some cases, the problem is completely equivalent to the canonical problem; in other cases certain transformations, approximations and simplifying assumptions are needed.

For instance, the split multiple pies game is an immediate representative for the problem of pirates dividing the bounty. However, it can also represent the negotiation over the price of a car through the following transformation. Let us consider the manufacturer's suggested retail price P_{MSRP} of the car and the dealer's invoice price $P_{invoice}$. In effect, the "pie" will be represented by the amount of money $P_{MSRP} - P_{invoice}$, which is the amount of profit split by the dealer and the buyer when negotiating a deal between them. Naturally, extended negotiations reduce this profit through inflation (which corresponds more or less exactly to the shrinking pie model), and/or through the cost of storage to the dealer, cost of rental car for the buyer and so on. These latter phenomena do not map directly in the canonical problem, but they can be approximated reasonably well by it.

There are, however, cases when the splitting multiple pies model, or its natural extensions can not capture the essential challenges of a class of real world problems (see Section 1.3). In this dissertation, we are considering the spatio-temporal negotiation problems. The issues under negotiation include actions such as meeting at certain locations at certain points in time, performing actions at certain locations before, at, or after specific time-points, or traversing certain paths with certain speeds. In this chapter, we propose an alternative canonical problem, the Children in the Rectangular Forest (CRF) problem, and we argue that (a) it represents many of the fundamental aspects of this class of problems and (b) it is simple enough to serve as the foundation of formal analysis.

In the following, we firstly describe the general convoy formation problem in Section 3.1. Then we propose the simplify version of it: the Children in the Rectangular Forest (CRF) problem in Section 3.2. We consider the CRF problem as the canonical problem representing the spatio-temporal negotiations. At the end of this chapter, in Section 3.3, we demonstrate the CRF problem matches well with the characteristics of spatio-temporal collaborations we discussed in 1.3.

3.1 The general convoy formation problem

Let us start by defining the convoy formation problem for embodied agents. Two agents A and B move from their source positions S_A and S_B to their destinations D_A and D_B. We assume that the agents move along the paths given by the function $P_A(t) \rightarrow L$, which we read by saying that agent A is at the location L at time t.

At the initial timepoint t_0 we have $P_A(t_0) = S_A$ and we define the *arrival time* of A as the smallest time t_{arr} for which $P_A(t_{arr}) = D_A$. For every path we define the *unit cost* $c_P(t)$, and the cost of a time segment $C(t_1, t_2) = \int_{t_1}^{t_2} c_P(t)dt$. Most of the time, we are interested in the cost of the path $C_P(t_0, t_{arr})$. In the simplest case we are only interested in the time to reach the destination. This corresponds to a unit cost $c_P(t) = 1$, and the cost of the path $C_P(t_0, t_{arr}) = t_{arr} - t_0$. Many environmental factors can be modeled by the appropriate setting of the unit costs. For instance, the unit cost might be dependent on the location $c_P(t) = f(P_A(t))$ or on the speed of the agent $c_P(t) = f(P'_A(t))$. Locations or speeds which are unfeasible to the agent can be set to have an infinite unit cost.

Two agents form a *convoy* if they are following the same path $P_{A+B}(t)$ over the period of time $[t_{join}, t_{split}]$. An agent is motivated to join a convoy because of the *convoy advantage*: the unit cost for the convoy is smaller than for the individual agent over the same path. One example is the case when convoys can traverse areas which are not accessible to individual agents: $\exists t \in [t_{join}, t_{split}] \;\; \exists l \;\; P_{A+B} = l$ with $c_{P,A}(t) = \infty$ and $c_{P,A+B}(t) = c \in \mathbb{R}$. Naturally, convoy and non-convoy segments of the path need to be continuous in space: $P_A(t_{join}) = P_B(t_{join}) = P_{A+B}(t_{join}) = L_{join}$ and $P_A(t_{split}) = P_B(t_{split}) = P_{A+B}(t_{split}) = L_{join}$. We call L_{join} and t_{join} the join locations and time, and L_{split} and t_{split} the split locations and time, respectively.

We are considering self-interested agents which are searching for the path with the smallest cost from source to destination. This path might or might not include segments traversed as a convoy. The agents should use negotiation to agree on the segment traversed as a convoy. The negotiation succeeds if an agreement is reached over a quadruplet $(L_{join}, t_{join}, L_{split}, t_{split})$. Convoy negotiation is thus a multi-issue negotiation, with two tem-

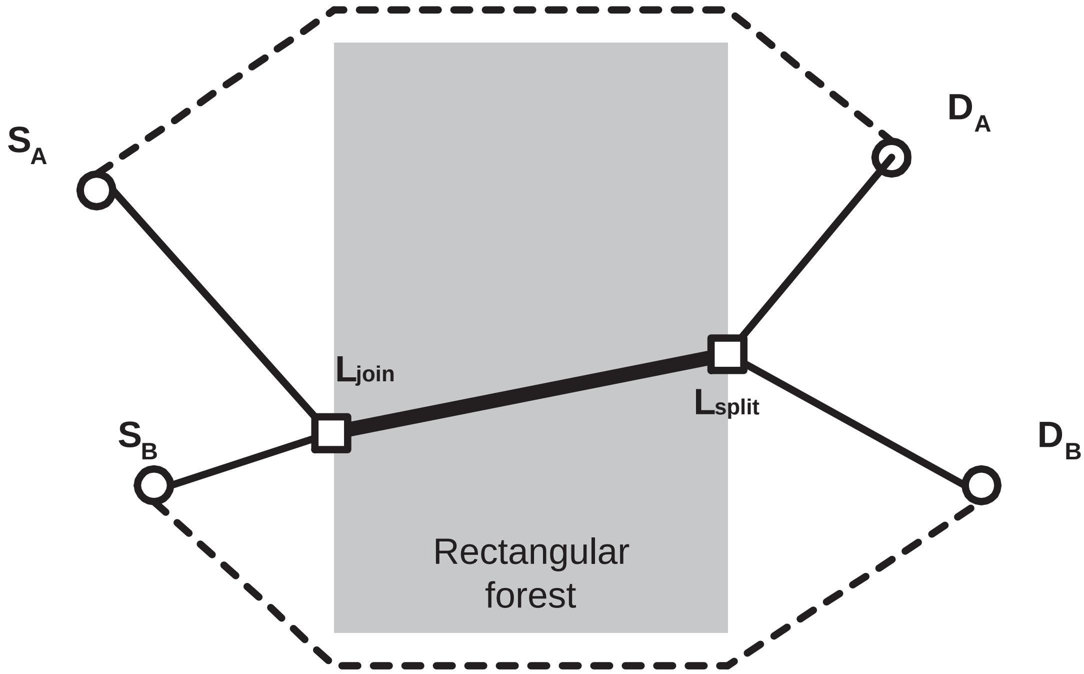


Figure 3.1: The Children in the Rectangular Forest problem. The trajectories associated with the conflict deal are shown with an interrupted line, while the trajectories corresponding to a possible agreement are shown with a continuous line.

poral and two spatial issues. It can be seen as a six-issue negotiation if we consider the spatial location $L = (x, y)$ as two issues.

3.2 The Children in the Rectangular Forest (CRF) Problem

In this section, we simplify the general convoy formation problem to the Children in the Rectangular Forest (CRF) problem. Let us assume that two children A and B are going from their departure locations S_A and S_B at one side of a rectangular forest of size $h \times w$, and they are going to their destinations D_A and D_B on the other side of the forest. The children were told not to go alone in the forest, but they can potentially traverse the forest together. The walking speed of the children (v_A and v_B) can be different, however when together, they will walk with the velocity of the slower child. The problem of the children is to use negotiation to agree on whether they will join up to traverse the forest together, and if yes, the join location, the join time, the point where they split and the time they split. If the negotiation fails, the conflict deal is a path which goes around the forest (see Figure 3.1).

We assume a rational behavior from the two children, that is, they will prefer the choice which takes them faster to their destinations. Let us consider some other properties of this problem.

Property 1. *The optimal trajectories of the conflict deal and the collaboration deal are a sequence of straight segments.*

The proof of this property is relying on elementary geometrical properties. What remains to be discussed is whether a rational agent would choose a curvilinear trajectory *during the negotiations* (note that the Property 1 only talks about the trajectories associated with the *deals*). The surprising answer is, yes. Let us consider an agent which might estimate the probability of a deal while waiting for an answer from a negotiation partner. An agent which is almost sure of a deal might move towards the predicted rendezvous point, while an agent which is almost sure of the conflict would move in the direction of the conflict deal trajectory. Between these two extremes, the agents might plan for an optimal trajectory, which strikes a balance between these choices. As the agents are moving during the negotiation time, the probability and utility of the deal changes in time. An optimal path therefore will be

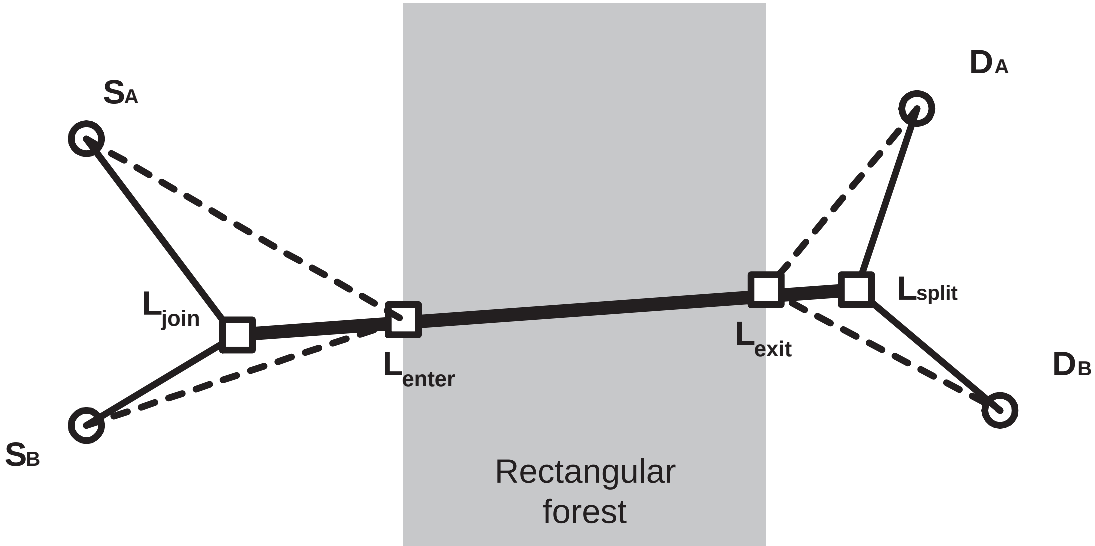


Figure 3.2: For any possible join and leave location, there is a join or leave location on the edge of the forest of equal or larger utility for both agents.

curvilinear, with edge points corresponding to events in the negotiation, such as the receiving of a new offer or the finishing of a utility calculation.

Property 2. *Deals where the join location is not on the edge of the forest are not Pareto optimal, or there is a deal where the join location is on the edge of the forest which provides the same utility to the agents.*

Proof: The proof of this property is a simple application of the triangle inequality. Let us make the assumption that there is deal $(L_{join}, t_{join}, L_{split}, t_{split})$, where the join location L_{join} is not on the edge of the forest (see Figure 3.2). Let us consider the point of entering the forest L_{enter} and the time to enter the forest t_{enter}. Then, from the triangle inequality:

$$dist(S_A, L_{join}) + dist(L_{join}, L_{enter}) \geq dist(S_A, L_{enter})$$

$$dist(S_B, L_{join}) + dist(L_{join}, L_{enter}) \geq dist(S_B, L_{enter})$$

That is we can build an offer $(L_{enter}, t_{enter}, L_{split}, t_{split})$, which is at least as good as the previous offer. In fact, if the strong inequality holds, we can build an offer which contains time values which are lower, that is, the offer as a whole has a higher utility.

$\square$.

Property 3. *Deals where the leave location is not on the edge of the forest are not Pareto optimal, or there is a deal where the leave location is on edge of the forest which provides the same utility to the agents.*

The proof of this property is analogous to the previous one (see Figure 3.2).

From general convoy formation point of view, the convoy advantage in the CRF problem is represented by the convoys ability to traverse a rectangular obstacle which is not accessible to the individual agents. The unit cost is either unity ($c_p(t) = 1$) or infinity ($c_p(t) = \infty$). Under these conditions, the cost of every feasible path is equal to the time to destination, but not all paths are feasible. The four negotiation issues are not completely independent in the CRF problem. For instance, if we know the maximum velocity of both agents, the split time t_{split} can be calculated from L_{join}, L_{split}, and t_{join}. Similarly, if all information is known about the current location and speed of the agents, the Pareto optimal value of t_{join} can be calculated, knowing L_{join}.

The CRF problem presents many challenges of the general convoy formation problem such as the difficulty of establishing whether an offer is feasible to the opponent, whether it represents a concession or not, and the difficulty of simultaneously negotiating temporal and spatial issues. At the same time, the CRF problem simplifies away the path planning problem, as all the Pareto-optimal deals correspond to paths formed of at most three linear segments.

3.3 Evaluating the CRF problem with characteristics of the spatio-temporal collaborations

In this section, we evaluate the proposed problem in the light of the five characteristics of the spatio-temporal collaboration problems we have highlighted in the Section 1.3.

(1) Heterogeneous issues

The CRF problem, as stated above, is a 4-issue negotiation, with two issues being points in the 2-dimensional space and two issues being time values. Depending of the assumptions of the problem, this can be farther simplified. For instance, if the velocities of both agents are known, the leave time is completely determined by the join and split locations and the join time, effectively reducing the problem to a 3-issue negotiation. By applying Properties 2 and 3, we can reduce the negotiation of the locations to a negotiation only on the y axis, another important simplifying factor.

The problem can be immediately extended to include a worth type issue, for instance by one of the agents offering some compensation to the other agent in exchange for a more favorable leave location.

(2) Non-monotonic valuation of issues

The issues under negotiation do not contribute linearly and monotonically to the utility of the agents. For instance, the join location and time has only an indirect impact on the time of arrival to destination, by its impact over what leave locations and times are feasible.

(3) Evolving world

The agents are moving during the negotiation, which makes the value of an offer dependent on the time at which it is evaluated and the state of the world. For instance, if an agent decided that an agreement is very likely, it moves towards the predicted join location, through this action increasing the value of the predicted deal. Alternatively, if an agent assumes that a deal is highly unlikely, it will move on the conflict deal trajectory, making the value of the offer lower and lower as it moves farther and farther away from the proposed join location.

(4) Offers need to be verified for feasibility

Not every offer is feasible in the CRF world, due to the limited velocities of the agents'. The feasibility conditions of an offer $(L_{join}, t_{join}, L_{split}, t_{split})$ when the current locations of the agents are L_A and L_B and the current time is t_{crt} are described by the following inequalities:

$$
\begin{aligned}
\frac{dist(L_A, L_{join})}{v_A} &\leq t_{join} - t_{crt} \\
\frac{dist(L_B, L_{join})}{v_B} &\leq t_{join} - t_{crt} \\
\frac{dist(L_{join}, L_{split})}{v_A} &\leq t_{split} - t_{join} \\
\frac{dist(L_{join}, L_{split})}{v_B} &\leq t_{split} - t_{join}
\end{aligned}
\tag{3.1}
$$

Note that in a scenario where the agents do not know the other agents location and velocity, the first and the third condition can be evaluated only by agent A, while the second and fourth condition only by agent B. We also note, however, that while the feasibility of an offer is described by multiple conditions and exhibits interesting variations depending on the amount of information disclosure, the calculations themselves are simple and computationally inexpensive.

(5) Interaction between negotiation time and physical time

The negotiation between two agents happens in the physical time of the movement. If a negotiation round i takes t_i time, the agents will move $v_A t_i$ and $v_B t_i$ respectively on their planned trajectories. How much each negotiation round takes depends on the algorithms deployed by the agents, the computational power of the agent whose turn it is, and the messaging overhead. Various scenarios can be modelled with relative ease (negotiation between a fast and a slow agent, negotiations on different physical scales etc).

While there is a rich set of modelling possibilities, the calculations are sufficiently simple to make both simulation based and (at least as long as straight segment based trajectories are considered) analytical approaches possible.

We conclude in this chapter that the Children in the Rectangular Forest (CRF) problem exhibits all the five characteristics of the class of problems of negotiating for spatio-temporal collaboration. The CRF problem can be seen as a simplified version of the general convoy formation problem. With some transformations, it can also serve as the model for the other problems mentioned before. By considering one of the agents to be the shuttle and the other agent the passenger it can model the problem of the transportation of elderly persons. One additional modification would be that in this case that the two agents will move with the velocity of the *faster* agent after joining. The case of soccer pass can be modeled by considering one of the agents to be the ball while the other agent being the player waiting for a pass. The common feature of these problems is that they are all dealing with negotiation about rendezvous at certain point and time - one possible name for these problems being *spatio-temporal coincidence problems*. Through immediate and natural extensions, many additional interesting behavior can be modeled. At the same time, significant simplifications can be applied to the calculations of utility and feasibility, and the model is sufficiently simple to make analytical study possible.

CHAPTER 4
EVALUATING THE SCENARIOS

Each of us has an intuitive feel for negotiation scenarios which are "easy" because the negotiation partners have a strong incentive to form a deal and for scenarios which are "hard" because a rational agreement is difficult to find (or it might not exist). Also, we have an intuition of certain negotiation scenarios where one of the participants has "more to gain" from an agreement.

Our objective in this chapter is to develop metrics which match well with these intuitions, while abstract away the other parameters of the scenario (such as the location and destination of the agents). Note that all the metrics we describe are based on the general convoy formation, but we also provide examples in the CRF problem. In the CRF problem, a scenario is defined by the map (the size of the forest), the source locations of the two agents S_A and S_B, their destination locations D_A and D_B, and their velocities v_A and v_B. The path of the agents are series of segments together with the velocities of the vehicle on the different segments.

4.1 The utility of the offer for individual agent

In the convoy formation problem, the unit cost is either unity ($c_p(t) = 1$) or infinity ($c_p(t) = \infty$). Under these conditions, the cost of every feasible path is equal to the time to destination, but not all paths are feasible. We will also assume that the agents negotiate in physical time, with each negotiation round taking time t_r.

We call the cost of an offer $C^{(A)}(O)$ of agent A for a particular offer $O = \{L_{join}, t_{join}, L_{split}, t_{split}\}$ the time it takes for the agent to reach its destination if it accepts the offer and follows the trajectory. The lower the time to destination, the more desirable is the offer for the agent. The time to destination is composed of three components: the time it takes for both agents to reach the meeting location, the time for traveling together during the convoy, and time from the split location to the agent's destination. We assume $C^{(A)}(O) = \infty$ if the offer is unfeasible for the agent A. The cost of the conflict deal $C^{(A)}_{conflict}$ is the time for the agent to reach its destination if it does not make any deal.

As time is passing during the negotiation, the actual cost of an offer made at negotiation round n will be $C^{(A)}_{r=n}(O) = n \cdot t_r + C^{(A)}(O)$. This also applies to the cost of the conflict deal at round n: $C^{(A)}_{conflict,r=n} = n \cdot t_r + C^{(A)}_{conflict}$.

Considering agents whose negotiation time is the physical time requires us to refine our definition of rationality of a deal. At the beginning of the negotiation, at time t_0, the agent has a conflict deal path with cost $C^{(A)}_{conflict}$. According to the *baseline rationality* definition, any offer which has a higher cost than $C^{(A)}_{conflict}$ is not rational and it will not be accepted by the agent. Any negotiation the agent might enter implies a risk of conflict. Thus, at negotiation round n the agent might find itself in the position that it has already incurred costs $C_x = n \cdot t_r$. If at this moment an offer with cost $C^{(A)}(O)$ is received, it will be called *pragmatically rational* if $C^{(A)}(O) < C^{(A)}_{conflict}$ and *baseline rational* if $C_{offer} + n \cdot t_r < C_{conflict}$. A rational agent will need to act based on the pragmatic rationality, as the original conflict deal alternative is not available any more at this moment in time. Occasionally, the agent might find it necessary to accept deals which are not baseline rational.

However, when we are measuring the overall performance of the negotiation strategy / action strategy pairs, the term of comparison should be the original conflict deal. In order for a strategy pair to be acceptable, it needs to be baseline rational at least in the statistical average.

Definition 1. *The* ***pragmatic utility of an offer O*** *for agent A, denoted with $U^{(A)}(O)$, is the time the agent saves accepting the offer compared to the conflict deal, considering no time spent on the negotiation.*

$$U^{(A)}(O) = C^{(A)}_{conflict} - C^{(A)}(O) \tag{4.1}$$

The ***baseline utility of the offer*** *which has been made at the n-th negotiation round is:*

$$U^{(A)}(O) = C^{(A)}_{conflict} - C^{(A)}(O) - n \cdot t_r \tag{4.2}$$

For instance, let us consider an agent whose time to destination is 45 minutes proceeding alone. Let us assume that the agent spent 15 minutes negotiating a deal which takes it to destination in 40 minutes. The pragmatic utility of this deal is +5 while the baseline utility is -10. At time 15, the negotiation time being already spent, the agent is better off taking the deal (which makes it arrive at time 55) than taking the conflict deal (which makes it arrive at time 60). Thus, the deal is pragmatically rational (at time 15). The deal however, is not baseline rational, because the original conflict deal was 45, thus the agent would have been better off if it does not negotiate at all.

In the CRF problem, for an offer $\mathbf{O} = (L_{join}, t_{join}, L_{split}, t_{split})$ made at time t_{crt}, the agent A with source location at S_A, current location at L_A and destination at D_A will arrive the destination at:

$$C^A(\mathbf{O}) = \begin{cases} +\infty \text{ if } t_{crt} + \frac{dist(L_A, L_{join})}{v_A} > t_{join} \\ +\infty \text{ if } \frac{dist(L_{join}, L_{split})}{v_A} > t_{split} - t_{join} \\ t_{split} + \frac{dist(L_{split}, D_A)}{v_A} \text{ otherwise} \end{cases} \tag{4.3}$$

The cost is assumed to be infinity if the offer is not feasible. Similarly we define the cost of the conflict deal as time spend in the negotiation until the current moment t_{crt}, plus the time necessary to reach the destination from the current location L_A by going around the forest. Note that the cost of both the collaboration and the conflict deal depend on the state (the current time and location of the agent). The implication is that the rationality of an offer is also state dependent. An offer might be rational for an agent at a certain moment in the negotiation, even if, together with the previous trajectory of the agent since the start of the negotiation it amounts to a total trajectory which is worse than the *original* conflict deal. The opposite case is also possible: an offer which would have been favorable at the beginning of the negotiation it might not be rational for the agent in the current state (for instance, if the agent is already well on its way towards the conflict deal).

4.2 Defining a collaborativeness metric

Metrics introduced in this section are calculated from the individual agent point of view. However, some of them are not accessible by the agents themselves, because the initial condition is zero-knowledge. As a result, metrics discussed in this section are served to analyze the scenario from the outside supervisor.

4.2.1 Metrics from individual agent point of view

Definition 2. *We define the* ***absolute best time to destination*** $C_{ab}^{(A)}$ *for agent* A *the time it would take it to reach the destination assuming an ideally performance and ideally collaborative negotiation partner.*

For the CRF problem, the trajectory associated to the absolute best time to destination is a straight line from the source to destination traversing the forest with the agent's maximum velocity.

$$C_{ab}^{(A)} = \frac{dist(S_A, D_A)}{v_A} \tag{4.4}$$

This assumes that there is an ideal negotiation partner, who is (a) willing to accept any geometric location for meeting and splitting points proposed by the agent, (b) its velocity is greater than or equal of the current agent and (c) its current position is such that it can reach the meeting point at a time earlier or equal with the time it takes agent A to reach it.

Note that for a practical scenario, the absolute best time to destination may not be feasible, even for an ideally cooperative negotiation partner.

Definition 3. *We define the* ***ability constrained best time to destination*** $C_{acb}^{(A),\{B\}}$*, of an agent* A *negotiating with an agent* B*, the time* A *can reach the destination assuming an ideally collaborative agent* B*.*

The ability constrained best time takes into account the physical limits of the negotiation partner and the scenario. The meeting and split point of the offer associated with the ability constrained best time might not be the one situated on the intersection of the straight line to destination with the forest. The offer(s) associated with $C_{acb}^{(A),\{B\}}$ might not be rational for agent B.

Let us consider an agent for which the absolute best deal involves meeting at point L_1 at time $t_1 = 20$, with the agent reaching its destination at time $t_{dest} = 100$. However, the opponent can not physically make it to the point L_1 in at time t_1, because it is too far away. We need to search for a deal which is feasible, for instance by extending the join time to $t'_1 = 30$. This would also extend the time to destination to $t'_{dest} = 110$. Alternatively, we can also modify the location of the join point.

Definition 4. *The* ***rationality constrained best time to destination*** $U_{rcb}^{(A),\{B\}}$ *for agent* A *negotiating with agent* B *is the time to destination of agent* A *which can be obtained assuming that agent* B *will accept any offer, as long as it is rational for* B*.*

For instance, let us consider a case the ability constrained best deal for the agent A would have a time to destination 100, with meeting at point L_1 and splitting at L_2. Let us assume that this trajectory is also feasible for agent B. It is still possible, however, that this trajectory would result for the B in a deal which is worse than going around the forest alone. One reason for this might be that the split point L_2 is too far from the B's destination D_B. B will not accept such a deal. A different deal would need to be negotiated, which, however, would normally be less advantageous for agent A.

As $C_{acb}^{(A),\{B\}}$ and $U_{rcb}^{(A),\{B\}}$ introduce successive restrictions over $C_{ab}^{(A)}$, we have:

$$C_{conflict}^{(A)} \geq C_{rcb}^{(A)\{B\}} \geq C_{acb}^{(A)\{B\}} \geq C_{ab}^{(A)} \tag{4.5}$$

Each of these time to destination values define a set of one or more concrete offers which actually achieve them. Thus we define a rationality constrained best offer of A to be an offer $O_{rcb}^{(A),\{B\}}$ such that

$$C^{(A)}\left(O_{rcb}^{(A),\{B\}}\right) = C_{rcb}^{(A)\{B\}} \tag{4.6}$$

4.2.2 Metrics from the social point of view

The metrics introduced until now characterize the scenario from the point of view of one of the agents. Let us now develop a metric which quantifies the desirability of a certain offer O from the point of view of the social good.

Definition 5. *We call the* ***social cost of the offer O*** *any function* $C_{social}(O) = C_{social}\big(C^{(A)}(O), C^{(B)}($ *which is monotonically increasing both with* $C^{(A)}$ *and with* $C^{(B)}$*:*

$$
\begin{aligned}
&\forall C^{(B)}, C_1^{(A)} \geq C_2^{(A)} \Rightarrow C_{social}(C_1^{(A)}, C^{(B)}) \geq C_{social}(C_2^{(A)}, C^{(B)}) \\
&\forall C^{(A)}, C_1^{(B)} \geq C_2^{(B)} \Rightarrow C_{social}(C^{(A)}, C_1^{(B)}) \geq C_{social}(C^{(A)}, C_2^{(B)})
\end{aligned}
\tag{4.7}
$$

We call denote with O_{social} the set of offers which minimize the social cost:

$$
O_{social} = \arg\min_{\mathbf{O}} \left(C_{social}(O) \right) \tag{4.8}
$$

Within the constraints of this definition, there are many possible functions which can serve as the social cost function. The choice of a specific function depends on the policy of the supervisor. One simple choice is to define the social cost as the sum of the individual costs.

$$
C_{social}(O) = C^{A+B}(O) = C^{(A)}(O) + C^{(B)}(O) \tag{4.9}
$$

Note however, that a social best offer might not be rational for both agents. We can define a rationality constrained social cost, which assumes a cost of plus infinity for the offers which are not rational for one of the agents:

$$
C_{rcsoc}(O) = \begin{cases} +\infty & \left(C^{(A)}(O) > C^{(A)}_{conflict}\right) \vee \left(C^{(B)}(O) > C^{(B)}_{conflict}\right) \\ C_{social}(O) & \text{otherwise} \end{cases} \tag{4.10}
$$

Based on this definition, we can define the set of rationality constrained social best offers O_{rcsoc} as:

$$O_{rcsoc} = \underset{O}{\operatorname{argmin}} \left(C_{rcsoc}(O) \right) \tag{4.11}$$

Definition 6. *We define as the* ***collaborativeness of the scenario*** *from the point of view of agent A, negotiating with agent B, the ratio of the utility of the rationality constrained social best deal to the maximum rationally obtainable utility:*

$$\Xi^{(A),\{B\}} = \frac{C^{(A)}_{conflict} - C^{(A),\{B\}}_{rcsoc}}{C^{(A)}_{conflict} - C^{(A),\{B\}}_{rcb}} \tag{4.12}$$

Let us verify that this definition satisfies our intuition about the collaborativeness of a scenario. In a fully competitive scenario, there is no rational deal possible, thus the cost of the rational deal will be the conflict deal, thus we have $\Xi^{(A),\{B\}} = 0$. On the other hand, we say that a scenario is fully cooperative from the point of view of agent A if the rationality constrained social best offer is also the rationality constrained best offer for agent A. In this case $\Xi^{(A),\{B\}} = 1$.

Definition 7. *We define the* ***relative utility of an offer*** *for agent A as the ratio of the utility of the offer to the maximum rationally obtainable utility:*

$$U^{(A),\{B\}}_{rel}(O) = \frac{C^{(A)}_{conflict} - C^{(A)}(O)}{C^{(A)}_{conflict} - C^{(A),\{B\}}_{rcb}} \tag{4.13}$$

The relative utility of the agent can range from 0 to 1. Notice that the relative utility of a deal does not tell us whether the agent has negotiated "better" than the negotiation partner. There are situations when both agents can reach the maximum relative utility.

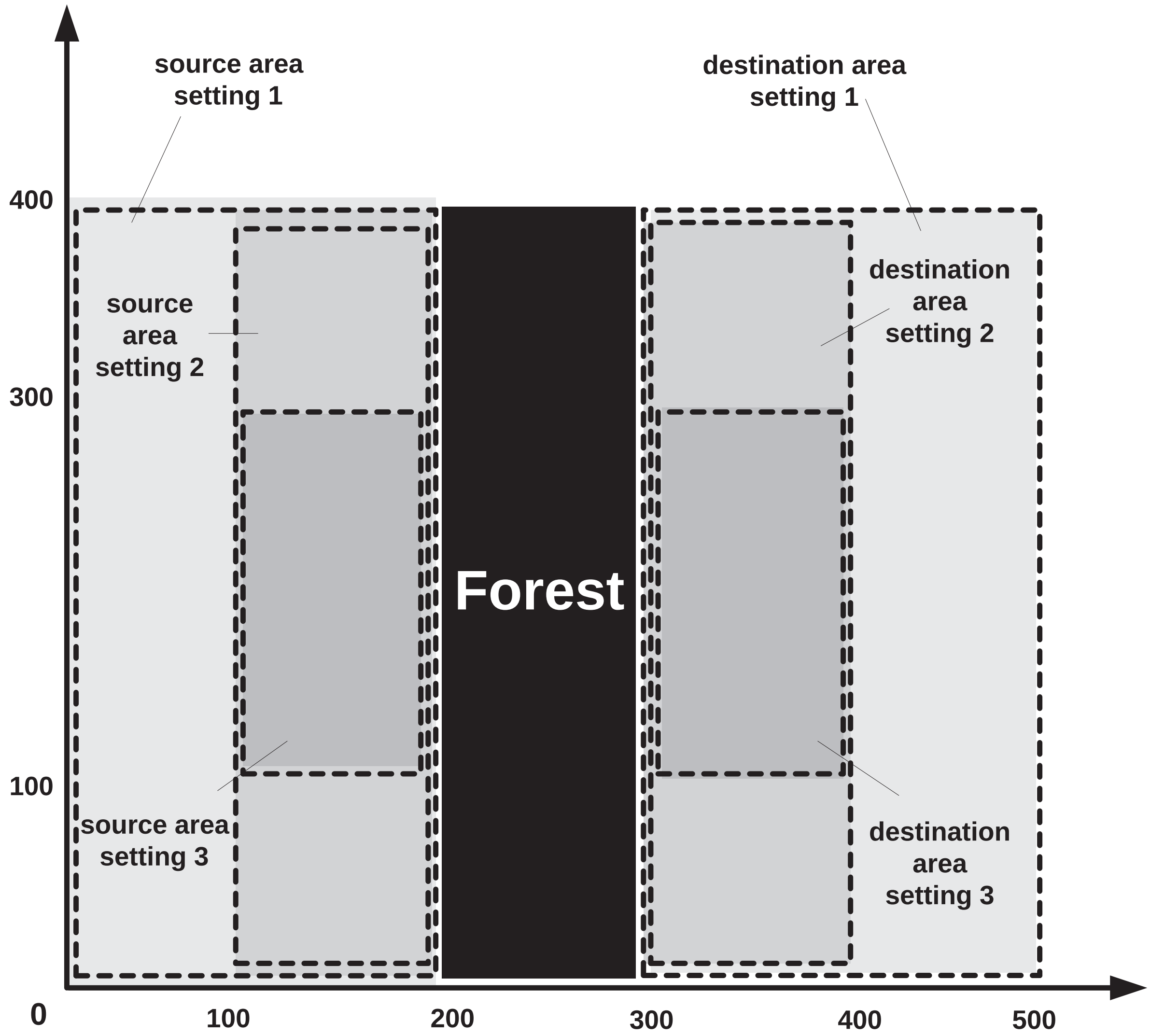
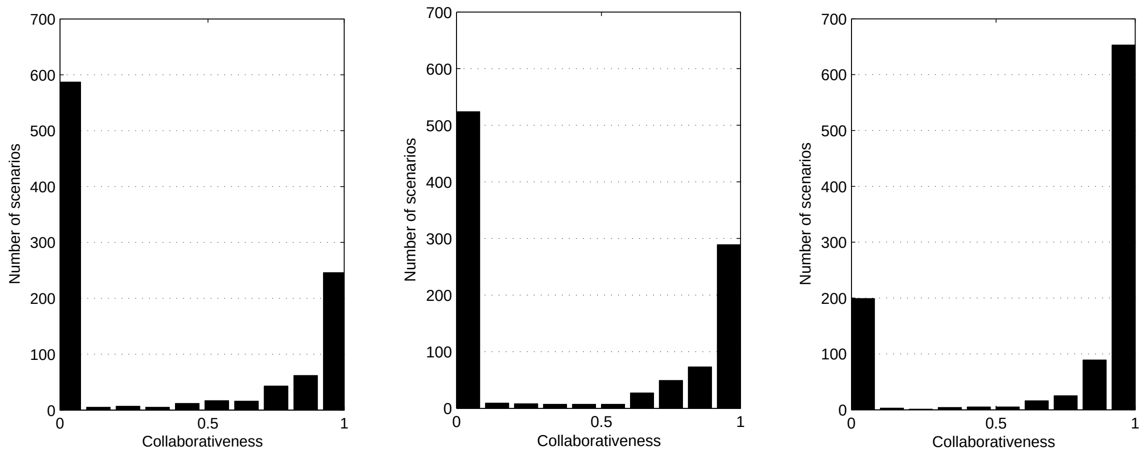


Figure 4.1: Three settings for the distribution of the source and destination areas for the study of the distribution of collaborativeness among scenarios.



(a) setting 1 (b) setting 2 (c) setting 3

Figure 4.2: The comparison of collaborativeness distributions in three cases of restricted areas.

4.3 Experimental study on the distribution of the collaborativeness

The distribution of the collaborativeness provides the answer to the following question: if we pick a random scenario, is it going to be competitive or collaborative? Naturally, the distribution of the collaborativeness depends on the distribution of the source and destination locations of the scenarios, as well as the distribution of the speed of the agents. Let us assume that the source and destination are distributed uniformly in rectangular areas situated immediately on the left and right side of the forest. To study a variety of possible distributions we consider three settings corresponding to the source and destination areas shown in Figure 4.1. For each setting, we generate 1000 scenarios by choosing the source and destination according to a uniform spatial distribution from the corresponding source and destination rectangles. We calculate the value of collaborativeness according to Equation

4.12, and assemble the values in a 10-bucket histogram. The three resulting histograms are shown in Figure 4.2.

We can make the following observations:

Setting 1: has the source and destination areas a square of the same height as the forest. The histogram shows a U-shape, with higher number of scenarios falling at the higher and lower extremes of collaborativeness.

Setting 2: has the source and destination areas rectangles of the same height as the forest but a width of half as much. The corresponding histogram shows a similar U-shape like in the previous case, but it is shifted towards the higher collaborativeness. We conclude that the closer is the forest to the source and destination, the higher the probability that forming a coalition to traverse the forest will be advantageous.

Setting 3: has the source and destination areas square and half the height of the forest. We find that the distribution of the collaborativeness is weighted toward the higher values.

This result matches our intuitive expectations. For instance, citizens in tightly packed cities such as New York and San Francisco rely more on public transportation, as their source and destination locations are frequently correlated. In cities spread over large areas such as Orlando or Phoenix, the transportation interests are rarely collaborative.

CHAPTER 5
NEGOTIATION PROTOCOLS

For most negotiation settings, it is assumed that the complexity of the negotiation relies on the strategy, while the protocol is a relatively trivial alternating exchange of offers by the two parties. Such a simple protocol would still work well for the CRF game with full knowledge. In the case of incomplete knowledge, however, the difficulty of forming a feasible offer as well as evaluating whether a given offer represents a concession or not, make simple offer-exchange protocols little better than random search. The simple protocol can be enhanced by schemes in which the agents add additional information of the negotiation flow to aid the negotiation partner in offer formation. In the following we illustrate the design space for the CRF negotiation protocols through several examples.

5.1 Simple exchange of binding offers (EBO)

In this simplest negotiation protocol, the agents are alternating in making fully specified offers in the form $\mathbf{O} = \{L_{join}, t_{join}, L_{split}, t_{split}\}$. The offers are binding for the agents who made the offer, in the sense that once made by an agent and accepted by the other agent, the offer will be the outcome of the negotiation. An example run of this protocol is illustrated in Figure 5.1-left.

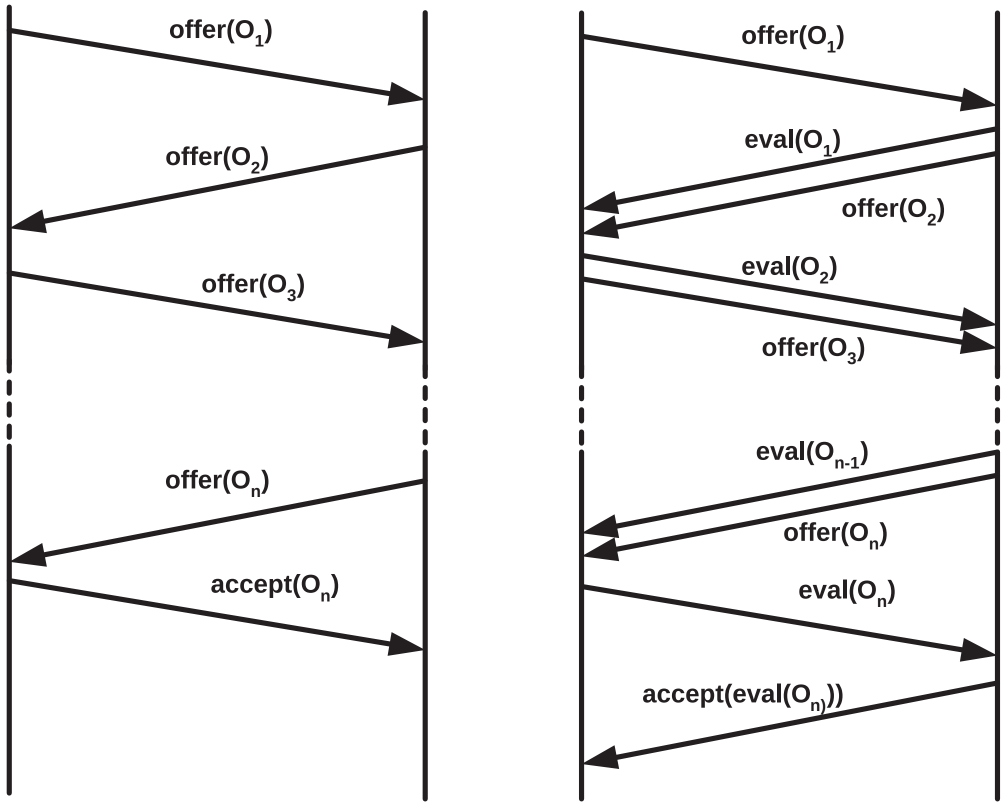


Figure 5.1: Example runs of negotiation protocols. (left) Exchange of binding offers (EBO). (right) Exchange of offers with mandatory, non-binding evaluations (EBOMNE).

5.2 Exchange of binding offers with mandatory, non-binding evaluations (EBOMNE).

In this protocol the agents are exchanging pairs of offers and evaluations. Agent A first chooses a spatially specified offer $O = \{L_{join}, ?, L_{split}, ?\}$, and computes the associated best time completion $BTC^{(A)}(O) = \{L_{join}, t^{(A)}_{join}, L_{split}, t^{(A)}_{split}\}$. This is the offer which A will send to agent B, which is guaranteed to be feasible for A and is binding for A. Agent B will calculate its own best time completion $BTC^{(B)}(O) = \{L_{join}, t^{(B)}_{join}, L_{split}, t^{(B)}_{split}\}$ for the same spatially specified offer. Using the two best time offers, B will form an *evaluation* of the initial offer

$$E(O) = \left\{L_{join}, \max(t^{(A)}_{join}, t^{(B)}_{join}), L_{split}, \max(t^{(A)}_{split}, t^{(B)}_{split})\right\} \tag{5.1}$$

This evaluation has the form of an offer which is feasible for both agents, but it is not binding for the evaluating agent. Rather it represents a *critique* of the original offer, and such it helps the other agent in the formation of feasible offers. Also, if the evaluation amounts to an offer which is not rational for the evaluating agent, an empty evaluation $\emptyset$ will be returned instead.

Let's see an example on how it works from individual agent point of view. At negotiation round i, agent A receives an offer O^B_{i-1} made by agent B at round $i-1$, the agent proceeds to *evaluate* it. If the offer is feasible and rational, the evaluation is the offer itself: $E^B_{i-1} = O^B_{i-1}$. If the offer is not feasible (for instance, because the agent can not reach the join location in time, or it can not match the required speed during the common path), the agent can extend the temporal components of the offer such that they become feasible for the agent. If the resulting offer is rational for the agent A, it will become the evaluation. If it is not rational, the evaluation is considered to be the empty set $E^B_{i-1} = \emptyset$. The evaluation will paired with a counter offer to form the return message. Thus the response of the agent A at negotiation round i will be the pair (O^A_i, E^B_{i-1}).

While the offers are binding, the evaluations are not. An empty evaluation intuitively means "the proposed spatial coordinates are very wrong", while an evaluation returned with a counteroffer means "I would be able to accept the offer, but I am not willing to". The evaluation does not immediately disclose the utility function of the agent, but they allow the opponent to select its offer more efficiently. Thus, the EBOMNE protocol represents a simple variant of *argumentation*.

At each round of this protocol, the agent can "accept" the opponent's offer, "confirm" the acceptance, "propose" a counter-offer, and "quit" the negotiation. When the agent choose propose a counter-offer, it would couple the evaluation of the received offer with the sending of new offer. An example run of this protocol is illustrated in Figure 5.1-right.

5.3 Exchange of binding offers with optional, non-binding evaluations (EBOONE).

A variation of the previous protocol removes the requirement that the agents evaluate every received offer. The advantage of this protocol is that agents would not be required to disclose information in response to offers which they would not consider. An agent would normally evaluate only offers which are satisfactory from the point of view of the spatial components.

Other combinations are also possible. For instance, exchange of offers with optional but binding evaluations is the (near) equivalent of a simple exchange of offers strategy where one of the agents is choosing as its next offer the evaluation of the opponents' offer.

There is an interdependence between the negotiation protocol and the strategies of the agents. A negotiation strategy created for the EBO protocol can be trivially extended to the EBOMNE as the evaluation can be created automatically - but the strategy would not take advantage of the information contained in the evaluations. The same strategy can be also trivially extended to EBOONE, by choosing not to send any evaluation. It is more difficult to "downgrade" strategies which rely on information from evaluations such as in EBOMNE protocol, to protocols where this information might not be available, such as EBO.

CHAPTER 6
NEGOTIATION STRATEGIES

In this chapter, we start to discuss the negotiation strategies used by the agents in the CRF problem. We separate strategies into two groups, according to their different negotiation protocols. The strategies in the first group are based on the simple Exchange of Binding Offers (EBO) protocol(introduced in Section 5.1). We found that the feasibility constraints are the main issues and the strategies on the EBO protocol are just a little better than the random search. However, we focus on offer formation process in these strategies. Offer formation means how agents generate the next counter offer and when they report no next offer or quit the negotiation. In the second group of strategies, all agents negotiate based on the protocol of Exchange of Binding Offer with Mandatory, Non-binding Evaluations (EBOMNE), introduced in Section 5.2. Agents are allowed to exchange arguments during the negotiation so that they get a better understanding to each other. Thus, they can easily propose mutually feasible and rational offers. At the end of this chapter, we compare the performance of these strategies in different scenarios with different collaborativeness.

It is worth to mention that all the strategies discussed in this chapter assume the agents are immobile during the negotiation. They also evaluate offers from the pragmatic utility point of view. We will allow agents to move while negotiating in Chapter 8 and discuss strategies which evaluate offers from baseline point of view in Chapter 7.

6.1 Strategies under the EBO protocol

Usually agents will start their negotiation by proposing an offer corresponding to its absolute best O_{ab}. Then, at each step, they have three options: "accept" the opponent's offer, "propose" another counter-offer or "quit" the negotiation with the conflict deal. Algorithm 1 describes the general strategy under the EBO protocol.

Algorithm 1 The general strategy under the EBO protocol

the agent receives an offer $\mathbf{O}_{opponent}$ from the opponent;
the agent calculates next offer $\mathbf{O}_{next}$;
if not exist O_{next} **then**
    if $\mathbf{O}_{opponent}$ is feasible and rational **then**
        accept $\mathbf{O}_{opponent}$;
    else
        quit the negotiation;
    end if
else
    if $U(\mathbf{O}_{opponent}) \geq U(O_{next})$ **then**
        accept $\mathbf{O}_{opponent}$;
    else
        propose the counter-offer O_{next};
    end if
end if

In this algorithm, when the agent receives an offer by the opponent, it firstly calculates the next counter offer. Then the agent tries to compare the utilities between its next offer and the offer proposed by the opponent. If the agent is in favor of the opponent offer, it will accept it and a deal is formed. Otherwise, it will propose the next offer and continue the negotiation. Note that the agent will always favor its own offer if the opponent's offer is either infeasible or irrational, as their utilities are negative values (minus infinity for the former case and negative value for the latter case). If the agent can not find the next offer (the agent tries all possible offers), it will make the final call: either accept the opponent's offer or quit the negotiation with the conflict deal. Next, we will discuss the different offer formation strategies.

6.1.1 Monotonic Concession in Space (MCS)

Although monotonic concession is one of the basic strategies in most negotiation settings, for the CRF game with incomplete information, monotonic concession is not possible. One compromise is to limit the concession to the spatial domain. This will usually, but not always represent a concession in terms of the utility function for the opponent.

The monotonic concession in space (MCS) agent is parametrized by the pair (c_m, c_s) representing the concession pace in the joining location and splitting location respectively. The agent will start its negotiation by proposing an offer corresponding to its absolute best $O^{(A),1} = O_{ab}^{(A)} = \{y_{join}^{(A),1}, t_{join}^{(A),1}, y_{split}^{(A),1}, t_{split}^{(A),1}\}$. The next offer of agent A is described by the following values:

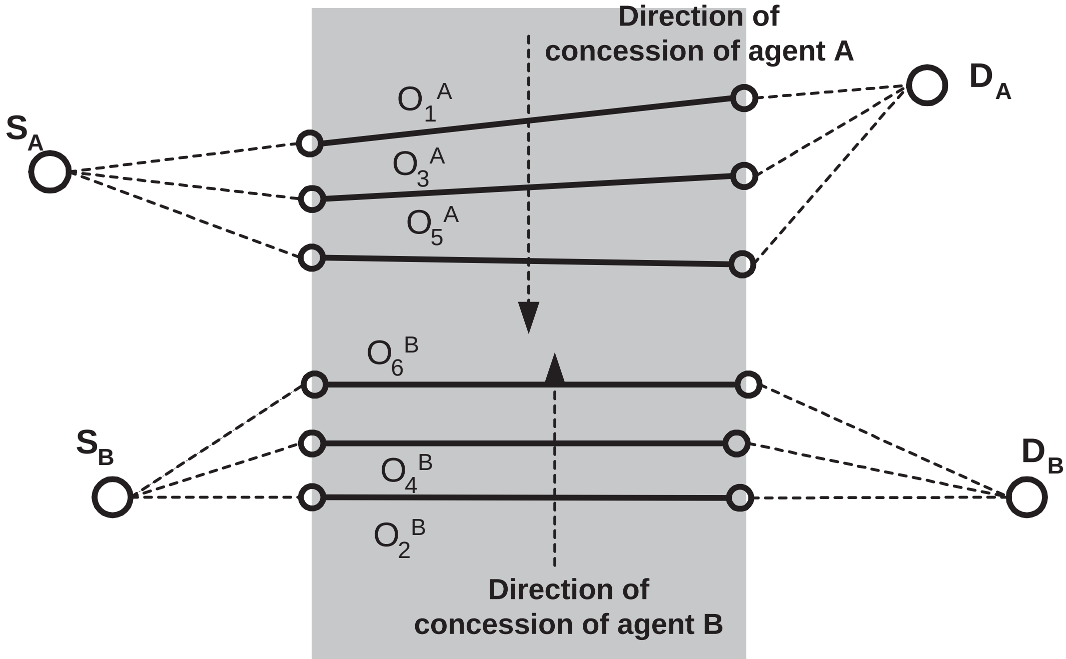


Figure 6.1: MCS agents propose their next offers in CRF game. The agent proposes its offers from absolute best and concede to the opponent's offer spatially with a fixed pace.

$$y_{join}^{(A),n+1} = \begin{cases} y_{join}^{(A),n} - c_m & \text{if } y_{join}^{(B),n} < y_{join}^{(A),n} - c_m \\ y_{join}^{(B),n} & \text{if } y_{join}^{(B),n} < y_{join}^{(A),n} \leq y_{join}^{(B),n} + c_m \\ y_{join}^{(B),n} & \text{if } y_{join}^{(B),n} > y_{join}^{(A),n} \geq y_{join}^{(B),n} - c_m \\ y_{join}^{(A),n} + c_m & \text{if } y_{join}^{(B),n} > y_{join}^{(A),n} + c_m \end{cases} \tag{6.1}$$

with a similar expression for $y_{split}^{(A),n+1}$. Using the resulting spatially specified offer $\{y_{join}^{(A),n+1}, ?, y_{split}^{(A),n+1}, ?\}$, the agent will calculate the best time completion as the next offer. Figure 6.1 illustrates the idea of this strategy. The agent monotonically concedes the spatial values to the opponent's current offer, and makes the final call when the two offers meet both in joining and splitting locations.

6.1.2 Internal Negotiation Deadline (IND)

One of the disadvantages for the MCS strategy is the agents have to decide the conceding pace at each side of the forest, and they don't know how many steps remained in the negotiation. In some scenarios where two absolute best offers are close in the space domain, the negotiation will end immediately.

In the internal negotiation deadline (IND) strategy, the agent sets to itself a deadline (expressed as a number of negotiation rounds) and adapts the speed of concession in function of the remaining rounds. If the number of rounds have expired without an agreement being reached, the agent breaks the negotiation with the final call.

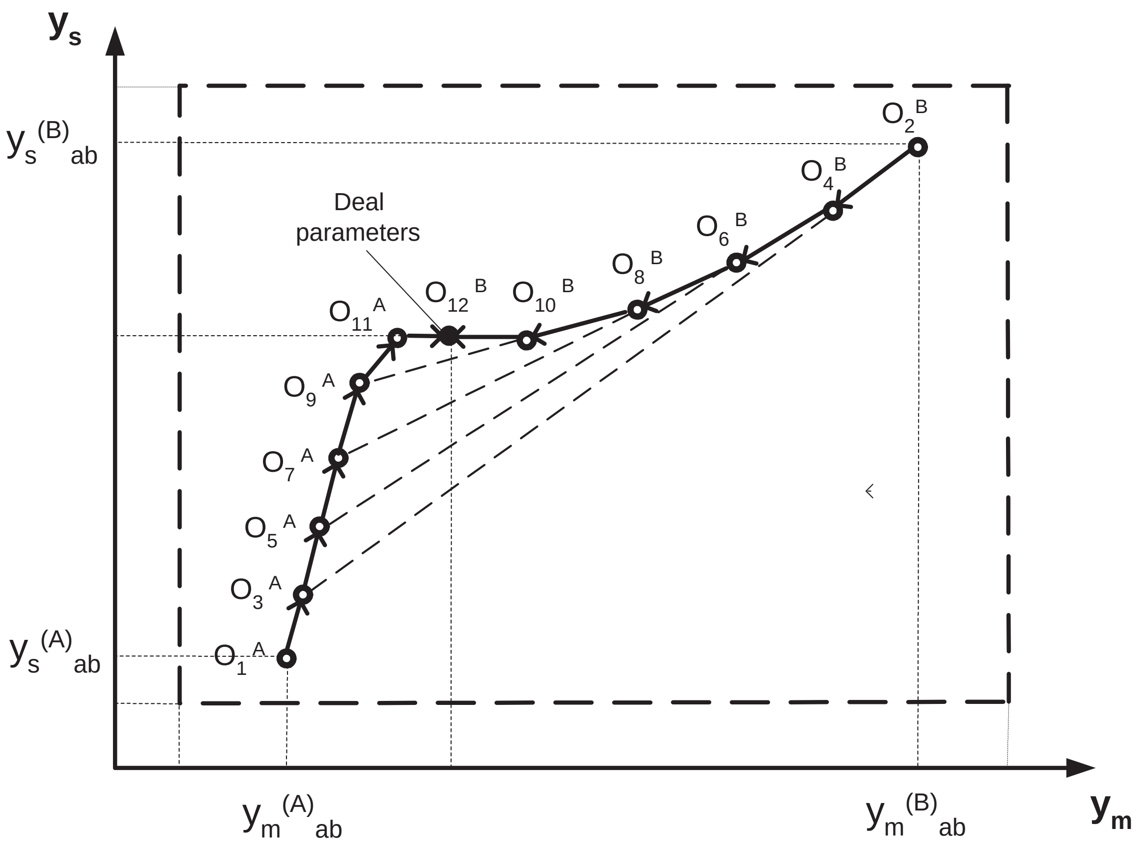


Figure 6.2: A negotiation trace between the a monotonic concession in space agent A and an internal negotiation deadline agent B.

Similar to the MCS strategy, the IND agent A starts by offering the absolute best $O_{ab}^{(A)}$. At every step the agent A will calculate the conceding distance based on the remaining rounds and the current spatial distance between two offers. The negotiation thus is guaranteed to be ended before the deadline. Naturally, a deal can be reached sooner if the opponent agent accepts an offer or its next offer is worse than the opponent's current one. Figure 6.2 shows a negotiation trace between a MCS agent and a IND agent. Note the adaptation of the concession speed by the IND agent.

6.1.3 Uniform Concession (UC)

The advantage of the IND strategy is it is easy to be understand and simple to be implemented. It resembles the monotonic concession strategy from single-issue worth-oriented domains. There are, however, some important differences. Conceding in the join and split location does not necessarily mean an even concession in terms of agent's own utility, nor a concession in opponent's utility. By exploring only specific combinations of joining and splitting locations, with the tight joining and splitting time according to the agent's own speed, the strategy excludes a large part of the solution space.

In the uniform concession strategy, the agent generates a pool of all possible offers (all combinations of joining and splitting location with a certain resolution). Also it adds possible time buffers at the meeting time field. The splitting time is calculated based on the minimum common speed in the history of all previous offers. The offer pool is then divided into a number of sub-pools. The first sub-pool contains offers which have the agent's absolute best utility $U_{ab}^{(A)}$. Each successive sub-pool $i = 1 \ldots n$ groups offers whose utility $U_{sp}(i)$ is smaller by the value α than the previous one, where $\alpha \in [0, 1]$ is the *conceding speed* of the agent:

$$U_{sp}(i) = (1 - (i \times \alpha)/2) \times U_{ab}^{(A)}, \text{for } (1 - (i \times \alpha)/2) \geq 0 \tag{6.2}$$

The insight is that from the agent's point of view all the offers in a given sub-pool are equivalent - however, for the opponent, the different offers in a sub-pool might provide different utilities. When conceding, the UC agent will simply pick the new offer from the next pool. Whenever the opponent's offer evaluates to a utility which is larger than the utility of the current sub-pool, the agent accepts the offer. Otherwise, it will calculate the next counter offer from the offer pool which is the most similar to opponent's offer O_{i-1}^{B}. The similarity between two offers is defined as the sum of squared difference for each issue:

$$\mathbf{O_i^A} = \arg\min_{\mathbf{O}}(||\mathbf{O} - \mathbf{E_{i-1}^B}||^2), \text{for } U^A(O) \geq U_{sp}(i) \text{ and } U^A(O) < U_{sp}(i-2) \tag{6.3}$$

If the agent reaches the last sub-pool without a deal, it quits the negotiation and takes the conflict deal. The offer formation in UC strategy is shown in Algorithm 2.

Algorithm 2 Offer formation for the Uniform Concession agent

Create $Set\langle \mathit{offer} \rangle$ to hold all possible offers;
while $Set\langle \mathit{offer} \rangle$ is empty **do**
  calculate $U_{sp}(i)$
  if $U_{sp}(i) \leq 0$ **then**
    return $O_{next} \leftarrow null$;
  end if
  find all *Offer* that $Utility(\mathit{Offer}) \in (U_{sp}(i), U_{sp}(i-2))$;
  add all *Offer* in $Set\langle offer \rangle$
end while
find the most similar *Offer* to $O_{opponent}$ in $Set\langle \mathit{offer} \rangle$;
return $O_{next} \leftarrow offer$;

6.2 Strategies under the EBOMNE protocol

The strategies discussed so far concern about how to compose the next counter offer and when to stop the negotiation with a final call. As we stated before, the most challenging job in the spatio-temporal negotiation is the difficulty to find mutual feasible offers, as agents do not know each other in zero-knowledge game. In this section, we introduce the strategies under the of Exchange of binding offers with mandatory, non-binding evaluations (EBOMNE) protocol. The idea is by adding arguments, agents can gradually understand each other. In different negotiation stage, agents have options to insist the previous offer and force the opponent to concede.

6.2.1 General ideas about how to use argumentation

Once the negotiation space is determined and the negotiation protocol agreed upon, the flow of the negotiation for a certain scenario is defined by the negotiation strategies of the agents.

The agents have a considerable freedom in choosing the negotiation strategy, which is limited only by the requirement of conformance with the protocol. However, the structure of the successful strategies is frequently dictated by the objective nature of the negotiation domain. In this subsection, we present some considerations about the stage of the negotiation, which need to be implicitly or explicitly made by any successful strategy. Let us consider that agent A has just received a message (O^B_{i-1}, E^A_{i-2}) from agent B. The agent A can evaluate the current negotiation stage as follows.

A *Blind search* $(E^A_{i-2} = \emptyset, U^{(A)}(E^B_{i-1}) < 0)$. In this case the agent A was told that its previous offer was not rational for B, but it also finds that the offer of the opponent is not rational for itself either. This situation frequently happens at the beginning of the negotiation. Being in this state does not necessarily means that there is no deal possible, but the agents need to explore the solution space for areas where mutually rational offers can be found.

B *Accept or concede* $(E^A_{i-2} = \emptyset, U^{(A)}(E^B_{i-1}) > 0)$: The agent's last offer (O^A_{i-2}) is not rational for the opponent but opponent's last offer (O^B_{i-1}) is pragmatically rational after extending the time issues. In this situation, the agent can either accept the opponent's offer with the time components extended or concede in a counter-offer. A deal will be formed only if the opponent confirms the modified offer.

C *Unbalanced blind search* $(U^{(A)}(E^A_{i-2}) < 0, U^{(A)}(E^B_{i-1}) < 0)$: The opponent returns an evaluation of the agent's last offer. However, this extended offer is not rational any more for the agent A. On the other hand, the opponent counter-offer is not rational for the agent, either. This situation can happen when the agent is near the forest while its opponent is not. Although the opponent accepts the joining and splitting locations, they fail to form a mutually rational agreement.

D *Opponent's offer acceptable* ($U^{(A)}(E^A_{i-2}) < 0$, $U^{(A)}(E^B_{i-1}) > 0$): The evaluation of the agent's previous offer was not rational for the agent, but the opponent's offer evaluation is rational. The agent can either accept the opponents offer, or create a counter-offer which it hopes to be rational to the opponent.

E *Agent's offer acceptable* ($U^{(A)}(E^A_{i-2}) > 0$, $U^{(A)}(E^B_{i-1}) < 0$): The evaluation of the agent's offer is pragmatically rational, while the opponent's offer is not. Intuitively, the agent has no motivation to concede until the opponent comes up with a rational offer. The agent will insist on its own offer until the opponent either accepts it, or provides a rational counter-offer.

F *Mutual concessions* ($U^{(A)}(E^A_{i-2}) > 0$, $U^{(A)}(E^B_{i-1}) > 0$): Both offers are evaluated to be rational, thus the agents now need to reach a deal with mutual concessions. Other things being equal, the agents will try to minimize their concessions. However, at the same time, the agents need to consider the risk of the opponent quitting the negotiation, as well as weight the potential benefits they can obtain from further negotiation against the time t_r lost in every negotiation round.

G *Accepted offer can not be confirmed* ($U^{(A)}(E^A_{i-2}) < 0$, $O^B_{i-1} = \emptyset$): The opponent accepted the evaluated version of the agent's counter offer. This evaluation, however, is not rational for the agent. As the opponent is also interested in getting to the split point as soon as possible, this means that no deal is possible with the current set of spatial components (it is not the matter of the opponent conceding more). The agent can either generate a spatially different counter-offer or quit the negotiation.

H *Accepted offer can be confirmed* ($U^{(A)}(E^A_{i-2}) > 0$, $O^B_{i-1} = \emptyset$): The opponent accepted the evaluated version of the offer, and this evaluation is rational for the agent. The agent can confirm the offer, which then becomes a deal. Alternatively, the agent can restart the negotiation with a new counter-offer if it considers that it can form the basis of a better deal.

6.2.2 Internal Negotiation Deadline with Argumentation(IND+A)

Similar to the IND strategy under the EBO protocol, the IND+A strategy sets up a deadline n_{max} (expressed as a number of negotiation rounds) and adapts the speed of concession in function of the remaining negotiation rounds. In negotiation stage E, however, the IND agent will insist its last offer to force the opponent to concede, because at that moment, the agent is sure that its previous offer saves opponent time while the opponent's offer does not. In the G and H stages, the IND agent stops calculating the next offer and makes decision between “quit” or “confirm”. In the other stages, the IND agent will calculate the next conceded offer described by the following values:

$$y_{join}^{(A),i} = \begin{cases} y_{join}^{(A),i-2} - c_m & \text{if } y_{join}^{(B),i-1} < y_{join}^{(A),i-2} \\ y_{join}^{(A),i-2} + c_m & \text{if } y_{join}^{(B),i-1} > y_{join}^{(A),i-2} \end{cases} \tag{6.4}$$

where the conceding amount in the meeting location is:

$$c_m = \frac{\left| y_{join}^{(B),i-1} - y_{join}^{(A),i-2} \right|}{\lceil (n_{max} - i)/2 \rceil}, \text{for } i < n_{max} - 2 \tag{6.5}$$

A similar expression for $y_{split}^{(A),i}$, the best time completion t_{join}^{A} and t_{split}^{A} are calculated accordingly. Note that there are three situations that the IND+A agent could not find the next offer: (a) the next concession break its own rationality constraint, (b) the current negotiation round i is greater or equal than $n_{max} - 2$ (one call left for the agent), and (c) the next conceded offer is worse than the evaluation of opponent’s previous offer. In these situations, the IND agent, again makes decision: either “accept” or “quit” the negotiation according to $U^{(A)}(E_{i-1}^{B})$. If the evaluation E_{i-1}^{B} is the same with the opponent’s offer O_{i-1}^{B} (no extension in time issues), the IND+A agent can “confirm” it directly.

6.2.3 Uniform Concession with Argumentation (UC+A)

In the same way to calculate the next offer as the UC strategy, the Uniform Concession with Argumentation (UC+A) agent exhaustively lists all possible offers, and divide them into a number of sub-pools. In stage A, B, C, D and F, the agent will keep on conceding (propose the offer which is the most similar with E^B_{i-1} in the next sub-pool) until the utility of opponent's offer is greater than the one in sub-pool. Same as IND+A agent, in stage E, the agent will hold the current sub-pool until the opponent concedes. In stage G, however, the agent keep on conceding by picking a random offer in next sub-pool as if the opponent did not accept its previous offer. In stage H, if the utility of extended offer $U^{(A)}(E^A_{i-2})$ is still less than the one in current sub-pool, the agent can continue the negotiation as if the agent did not accept its previous offer. Note that this is different with IND+A, because the IND+A strategy relies on the opponent's offer to generate the next counter offer, while the UC+A strategy has its own level of utility which is stored in the memory. The detailed algorithm for UC+A strategy is shown in Algorithm 3.

6.3 Experimental study on the performance of the strategies

In this section, we study the relative utility achieved by the specific strategies against each other under scenarios with various levels of collaborativeness. Naturally, if a deal fails through, the relative utility is zero.

One of the difficulties of our study is that we cannot generate directly random scenarios with predefined collaborativeness level. Thus we rely on *rejection sampling*, a technique borrowed from Monte Carlo simulation methods: we generate scenarios by picking source and destination points according to a uniform distribution, calculate their collaborativeness,

Algorithm 3 the Uniform Concession with argumentation strategy

```
agent $A$ received a message from agent $B$;
if the message is $(O^B_{i-1}, \emptyset)$ then
   $E^B_{i-1} \leftarrow evaluate(O^B_{i-1})$;
   calculate next offer $O^A_i$;
   if $O^A_i == null$ then
      return final call;
   else if $U^{(A)}(E^B_{i-1}) \geq U^{(A)}(O^A_i)$ then
      return accept or confirm $E^B_{i-1}$;
   else
      if $U^{(A)}(E^B_{i-1}) \geq 0$ then
         return propose $(O^A_i, E^B_{i-1})$;
      else
         return propose $(O^A_i, \emptyset)$;
      end if
   end if
else if the message is $(O^B_{i-1}, E^A_{i-2})$ then
   $E^B_{i-1} \leftarrow evaluate(O^B_{i-1})$;
   calculate next offer $O^A_i$;
   if $O^A_i == null$ then
      return final call;
   else if $U^{(A)}(E^B_{i-1}) \geq U^{(A)}(O^A_i)$ then
      return accept or confirm $E^B_{i-1}$;
   else if $U^{(A)}(E^A_{i-2}) > 0$ and $U^{(A)}(E^B_{i-1}) < 0$ then
      return insist $O^A_{i-2}$;
   else
      if $U^{(A)}(E^B_{i-1}) \geq 0$ then
         return propose $(O^A_i, E^B_{i-1})$;
      else
         return propose $(O^A_i, \emptyset)$;
      end if
   end if
else if the message is $(\emptyset, E^A_{i-2})$ then
   if $U^{(A)}(E^A_{i-2}) > U_{sp}(i)$ then
      return confirm $E^A_{i-2}$;
   else
      get random $O^A_i$ in the subpool;
      if $O^A_i == null$ then
         return quit;
      else
         return propose $(O^A_i, \emptyset)$;
      end if
   end if
end if
```

group them into 20 buckets, and then randomly reject scenarios from the buckets which are too full until they are uniformly filled with 500 scenarios each. The resulting collection of 10,000 scenarios was used in this experiment. Once the negotiation terminated (either with an agreement or a conflict), we measured the relative utility of the deal.

To present both the variability in the results of negotiation, as well as the underlying trends, we choose to superimpose a scatter-plot of the simulation results with a plot of the average values calculated on a per-bucket basis. Plotting only the average value would be misleading, as the spread of the simulation values is not accidental, but an intrinsic property which would not disappear if we would, for instance, run a larger number of experiments.

Figure 6.3 shows the scatter-plot and average values of relative utility for the case of two MCS agents negotiating against each other. We find that, as expected, the relative utility is increasing with the collaborativeness of the scenario. However, the scatter-plot shows that the results were spread over a large range of relative utilities. Even for scenarios with very high collaborativeness levels, there are many negotiations which end without an agreement. In general, this reflects weaknesses of the negotiation strategy under the EBO protocol, as the agents fail to get mutual feasible offers, and the negotiation breaks down while there were possible deals which would have been mutually acceptable.

The second experiment compares the internal negotiation deadline with and without argumentation (IND and IND+A). The results are shown in Figure 6.4. The shape of the IND vs. IND is roughly similar to the MCS vs. MCS graph, however, adding argumentation shows a great improvement. For instance, for a collaborativeness of 0.075 the IND+A vs.

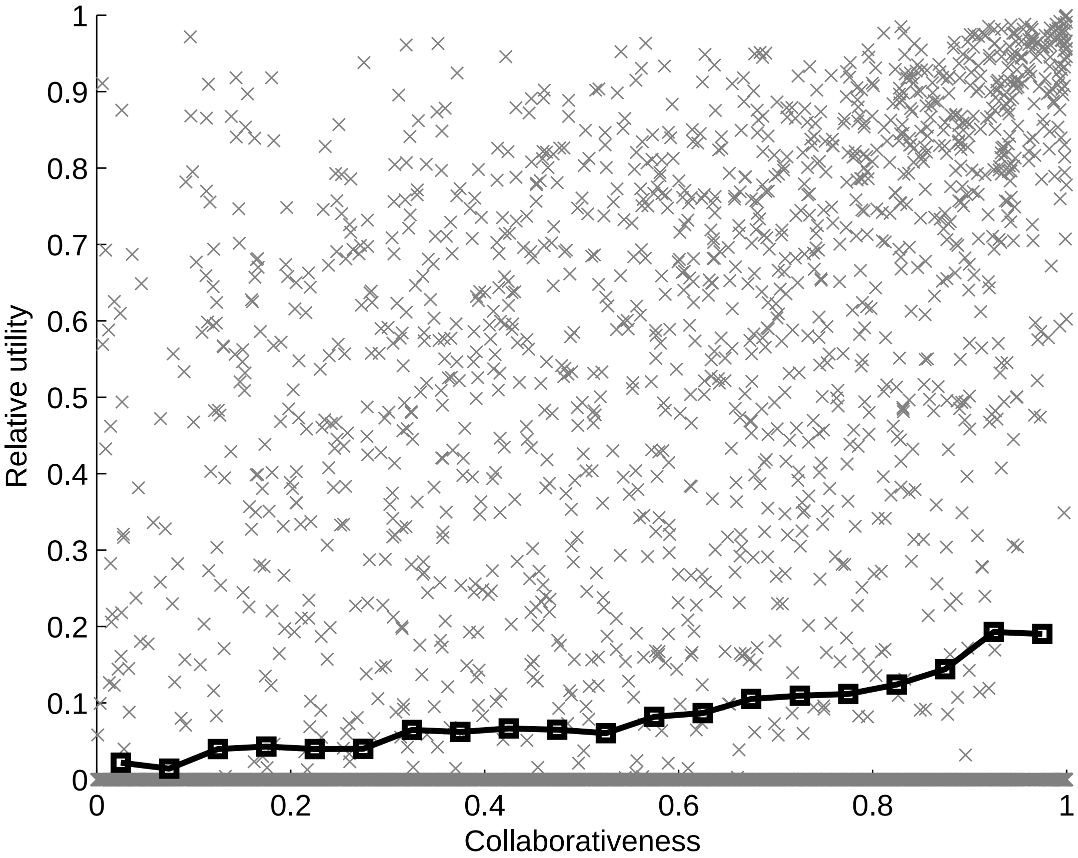
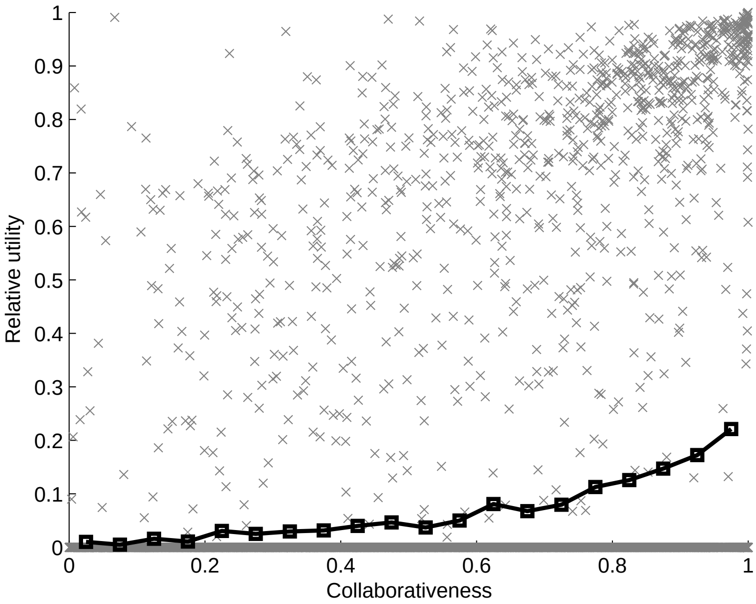
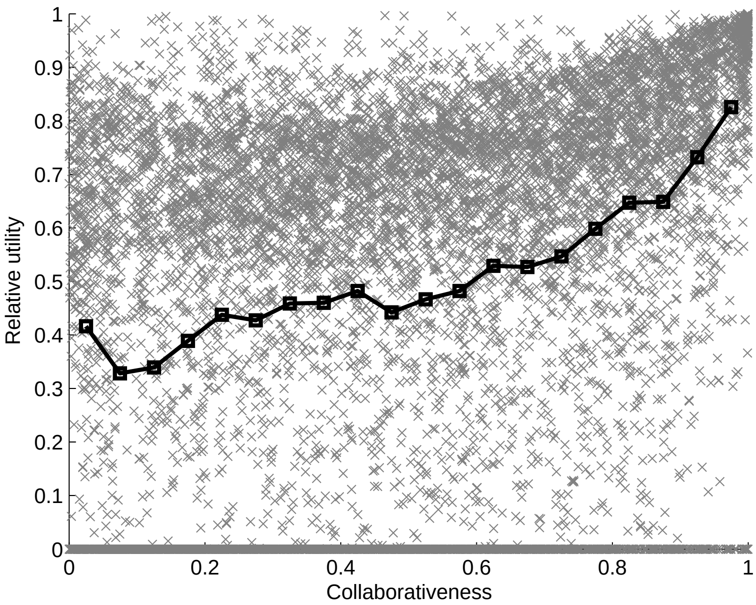


Figure 6.3: The scatter-plot and average values function of collaborativeness for two monotonic conceding in space (MCS) agents with identical parameters negotiating with each other.



(a) IND vs IND



(b) IND+A vs IND+A

Figure 6.4: The scatterplot and average values function of collaborativeness for two internal negotiation deadline (IND) and two internal negotiation deadline with argumentation (IND+A) agents with identical parameters negotiating with each other.

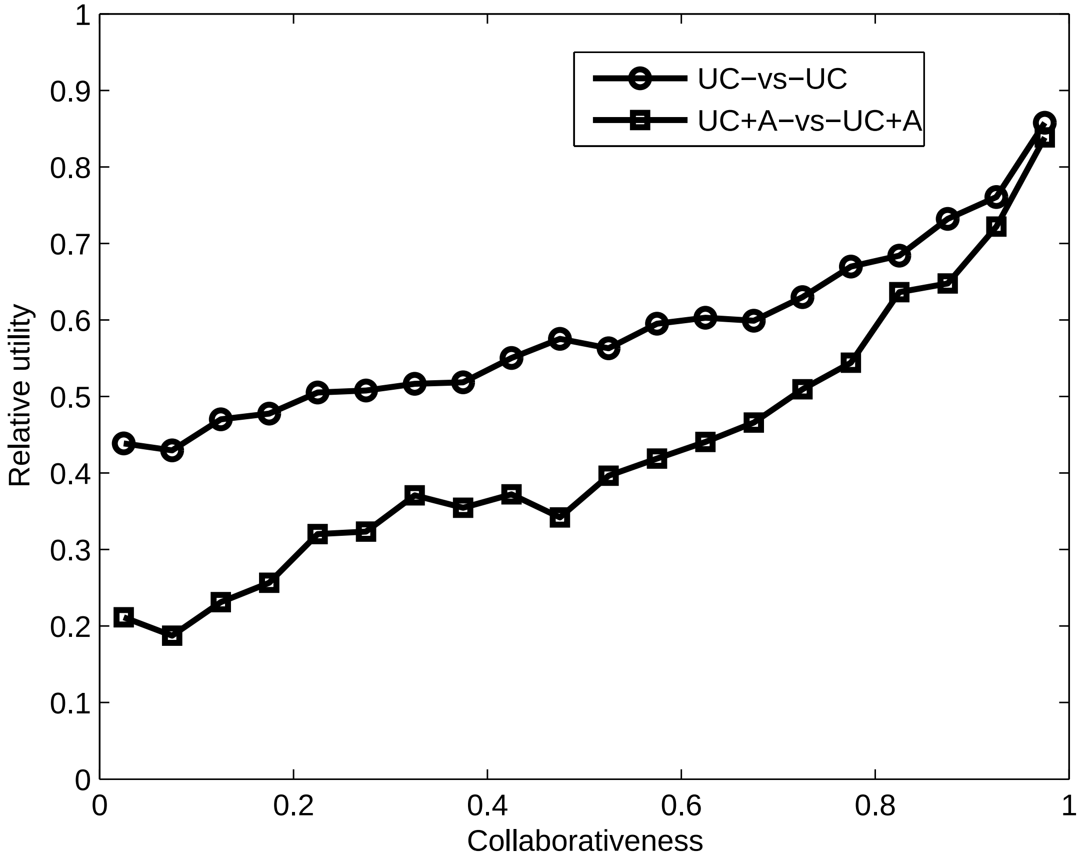


Figure 6.5: The averaged relative utilities in the function of collaborativeness for the Uniform Concession with and without argumentation (UC and UC+A) agents with identical parameters negotiating with each other.

IND+A settings obtains an average relative utility of 0.328, compared to 0.003 for the IND vs. IND case, and we can see a lot of successful deals are formed when agents exchange argumentation.

The third experiment compares the uniform concession with and without argumentation (UC and UC+A) strategies. Figure 6.5 shows the average relative utilities for UC vs. UC and UC+A vs. UC+A. We find that, even it doesn't use argumentation, the UC strategy gains much more relative utilities than the MCS and the IND. That also demonstrates searching solution space in the utility point of view is better than the concession in spatial domain. Then, we find that adding argumentation in strategy UC+A, however, decreases the relative utility comparing UC vs. UC. The reason is complicated. Comparing the UC strategy, UC+A strategy enables agents to insist their offers according to the the current stage of the negotiation. This means the UC+A agents try to balance each other, and find out a fair solution in both sides. On the other hand, according to the definition of collaborativeness, the social best solution is the one which saves most cost from the supervisor point of view. If the agent gains the same relative utility as the collaborativeness of the scenario, it turns out to be a wise agent who can think in both side of negotiation partner. The figure tells us when UC+A agent negotiates with each other, they are less greedy and the relative utility is more closer to the collaborativeness of the scenario which is the diagnoal in the figure.

In the previous examples, we always had the same types of agents negotiating each other. Pitting agents of different types against each other can give as insight into their relative performance. Figure 6.6 shows the average relative utility in the function of collaborativeness when an IND+A agent negotiates with a UC+A agent. The two plots represent the results for the same series of experiments seen from the point of view of the two agents (that is the $U_{rel}^{(IND+A),\{UC+A\}}$ and the $U_{rel}^{(UC+A),\{IND+A\}}$ values). For reference, we also added the plot of two IND+A agents and two UC+A agents negotiating each other. The overall shape of the curves is what we expected, the relative utility increases with the collaborativeness. However, the agent using the IND+A strategy consistently achieves higher utility values

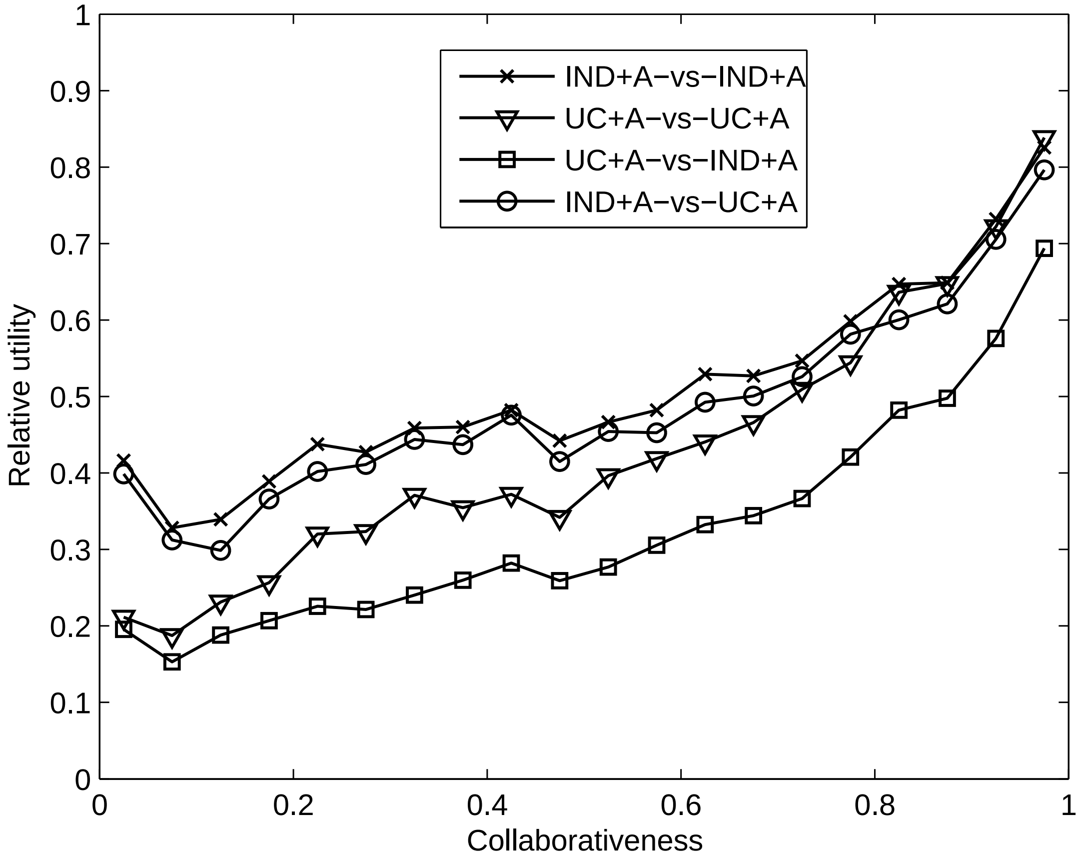


Figure 6.6: The comparision of average relative utilities both agents get when IND+A agent negotiates with the UC+A agent in different scenarios.

than the UC+A agent, which shows the IND+A strategy is more greedy than the UC+A strategy.

An interesting observation can be made comparing the results of the IND+A agent negotiating the different opponent. The utility values when the opponent is the UC+A agent are always a little lower than the case when the opponent is IND+A agent. That is because UC+A strategy has ability to restart the negotiation while the IND+A makes the final call. When facing the same opponent, the IND+A strategy performs better against the UC+A strategy in the low collaborativeness levels. It means the efforts of UC+A to find better deals improves the deals for the IND+A agent. For the high collaborativeness levels, however, the situation is reversed: for these scenarios the two agents have largely aligned interests, the UC+A can search the solution space more thoroughly to find better deals for itself.

CHAPTER 7
LEARNING WHILE NEGOTIATING

In this chapter, we investigate some sophisticated strategies which allow agents to learn the opponent during the negotiation.

The strategies we discussed in the previous chapter evaluate offers by their pragmatic utilities. That is, the agents assume the negotiation does not take time and they compare the utility of the offer with the cost of current conflict deal. However, as the negotiation in the CRF problem happens in the physical time, at the end of negotiation, the agents may form irrational agreement from baseline point of view. This kind of phenomenon is more significant in the scenarios which have small collaborativeness. In Section 7.1, we describe how strategies under the EBOMNE protocol can be augmented by collaborativeness analysis: we approximate the collaborativeness metric in the first several negotiation rounds, and use the result to cut short the negotiation when the estimated collaborativeness is lower than a threshold. Through experimental studies, we show that augmenting the strategies with collaborativeness analysis significantly improves their performance for low collaborative scenarios, and encounter only a minimal penalty in high collaborative scenarios.

On the other hand, the strategies under the EBO protocol suffer in finding mutual feasible offers. In Section 7.2, we let agent learn the opponent to improve the performance. We assume that the agents do not disclose any of their information voluntarily: the learning needs to rely on the study of the offers exchanged during normal negotiation. The opponent model is represented by the physical characteristics of the agent: the source location and the destination. The learning agent uses Bayesian learning to guess the opponent's model and update the beliefs dynamically during the negotiation. Furthermore, the belief about opponent's model can be contributed to re-construct the opponent's utility function and thus provide more acceptable offers to the opponent in the next round. Through experiments we show that agents which act according to learning strategies outperform a large set of benchmark strategies with no negotiation-time learning.

7.1 Augmenting strategies with collaborative analysis

In our current setting, negotiation happens in physical time, each negotiation round taking time t_r. The decision to close the negotiation (by either accepting the current offer, or by quitting with the conflict deal), should depend on the agent's view of the possible benefits it can obtain if it continues the negotiation weighted against the time delays this would involve. If no deal is possible, the agent is wasting utility by negotiating.

The collaborativeness metric we introduced in Equation 4.12 was developed precisely for the purpose of characterizing the potential deals in a scenario. We need to emphasize that a high collaborativeness metric does not necessarily guarantee a negotiation success, because the agents need to find those mutually beneficial deals, which depends on the offer formation strategies. On the other hand, even if the collaborativeness is low, the agent might hope to "trick" the opponent in a deal which is only marginally rational for the opponent, but much better for the agent. Overall, however, the collaborativeness metric is a good predictor of negotiation success of certain scenario.

Thus, it makes sense to augment the negotiation strategies with collaborativeness analysis. These augmented strategies would alter their behavior in function of the collaborativeness of the current scenario, for instance, quitting earlier the negotiation for low collaborativeness scenarios.

7.1.1 Estimate the collaborativeness and adapt conceding speed

The problem with the collaborativeness metric $\Xi^{(A)}$ is that it can be evaluated only by a full knowledge agent (e.g. a supervisor). The agent participating in a negotiation starts with zero knowledge, but it can gradually acquire information from the negotiation. At the second round of the negotiation, the agent assumes its absolute best time $C_{ab}^{(A)}$ as the rationality constrained best time $C_{rcb}^{(A)\{B\}}$. It will approximate the rationality constrained social best time $C_{rcsoc}^{(A),\{B\}}$ as the averaged utility between its first offer and evaluation of the opponent's first offer. Thus the agent will estimate the collaborativeness as:

$$\Xi \approx \Xi_{estimate}^{(A)}(E_2^B) = \frac{C_{conflict}^{(A)} - \frac{C^{(A)}(O_1^{(A)}) + C^{(A)}(E_2^B)}{2}}{C_{conflict}^{(A)} - C^{(A)}(O_1^{(A)})} = \frac{U_{ab}^{(A)} + U^{(A)}(E_2^B)}{2 \times U_{ab}^{(A)}} \tag{7.1}$$

If this value is negative, it can be viewed as non-collaborative scenario with collaborativeness of zero.

Let us now see how this value can be used by the negotiation strategies. The internal negotiation deadline with argumentation augmented with collaborativeness analysis (IND+A+CA) will compare the estimated collaborativeness with a threshold $\Xi_{threshold}$. If the estimate is smaller, the agent will quit the negotiation either by accepting the opponents first offer (if it is rational) or by taking the conflict deal. For $\Xi^{(A)}_{estimate}(E^B_2) > \Xi_{threshold}$ the IND+A+CA agent will change its negotiation deadline according to the following formula:

$$n'_{max} = \begin{cases} 0 & \text{if } \Xi_{estimate} < \Xi_{threshold} \\ n_{max} \times \frac{\Xi_{estimate} - \Xi_{threshold}}{1 - \Xi_{threshold}} & \text{otherwise} \end{cases} \tag{7.2}$$

The uniform concession with argumentation and collaborativeness analysis (UC+A+CA) agent, will also compare the estimated collaborativeness with a threshold $\Xi_{threshold}$. If the estimate is smaller, the agent will quit the negotiation either by accepting the opponents first offer (if it is rational) or by taking the conflict deal. For $\Xi^{(A)}_{estimate}(E^B_2) > \Xi_{threshold}$ the IND+A+CA agent will change its conceding pace α according to the following formula:

$$\alpha' = \begin{cases} 1 & \text{if } \Xi_{estimate} < \Xi_{threshold} \\ \alpha \times \frac{1 - \Xi_{threshold}}{\Xi_{estimate} - \Xi_{threshold}} & \text{otherwise} \end{cases} \tag{7.3}$$

The intuition behind states that the absolute best offer can get the agent to reach its destination, assuming it has an ideal opponent in an ideal scenario. Its utility should be similar with the rationality constrained best utility, as the latter one just adds two geometric restrictions in the solution space. On the other hand, taking the average between the utilities of the first two offers seems to be fair, if the utility function is linear and both agents concede until they meet in the middle. In a low collaborative scenario, the two agent need long time to search the deal. Even they form an agreement at last, the utility of the deal may not compensate the cost of negotiation. In this case, it should be wise to drop the negotiation immediately or increase the conceding pace so that they can end the negotiation quickly.

The estimation above did not consider the impact of time scale t_r. If each negotiation round takes too much time, the agent should continue to accelerate the process of negotiation. We let the agent remember the best evaluation in history which has the most pragmatical utility, and continuously check if such evaluation is irrational from baseline point of view. If the agent finds out the its baseline utility is less than a threshold, it will drop the negotiation by either sending "accept" message or "quit" the negotiation directly. Let's talk about the intuition behind this. When the negotiation time t_r is expensive, the baseline utilities of all un-explored deals are decreasing quickly. The best potential deal which has already been explored by both agents somehow indicates this decreasing speed. If its baseline utility is less than a threshold, the agent should quit the negotiation immediately to avoid the further damage.

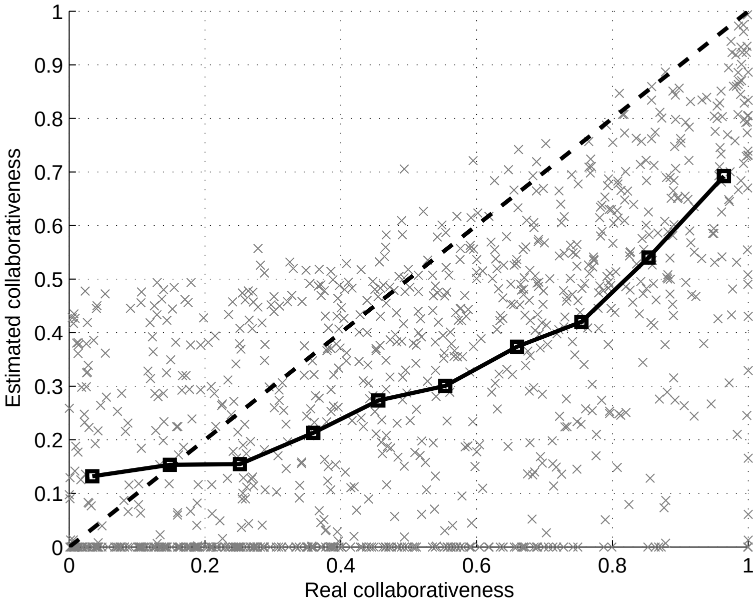


Figure 7.1: The estimated vs. real collaborativeness. Each point in the scatter plot corresponds to one scenario. The solid line is the average estimate for different collaborativeness values.

7.1.2 Experimental study on augmenting strategies

In this subsection, we focus on comparing the strategies with collaborativeness analysis versus the one without it.

7.1.2.1 Accuracy of collaborativeness estimation

Figure 7.1 shows the scatter plot and the average of the estimated collaborativeness $\Xi_{estimate}$ function of the real value of the collaborativeness. In this graph, every point represents the estimate at negotiation round 2 for a total of 1000 scenarios. The closer it is the point to the diagonal, the better the estimate. The first observation is that the estimate is by no means perfect. Quite a number of datapoints fall far from the diagonal. There are even cases where a fully collaborative scenario is estimated to have near-zero collaborativeness. There are, however, no cases where low collaborativeness scenarios are estimated to have high collaborativeness. The average value, on the other hand, is tracking the diagonal relatively well, although it is always below the diagonal. Agents using this metric will likely err on the side of safety, underestimating collaborativeness rather than overestimating it.

Overall, the estimate of collaborativeness is satisfactory, considering that we are only two negotiation rounds in a negotiation started with zero knowledge. It also opens the possibility of future work towards of a more accurate estimation based on information acquired in subsequent negotiation rounds.

7.1.2.2 Negotiation performance for augmenting strategies

In the following we investigate the performance of the negotiation strategies IND+A and UC+A and their variants augmented with collaborativeness analysis IND+A+CA and UC+A+CA. In a setting where the negotiation takes place physical time, the negotiation performance of the agents can be considered from two points of view. The pragmatic relative utility (Definition 7) measures the balance between the negotiation results of the participating agents. We already see these values in Section 6.3, and this is a pragmatic measure which does not depend on the negotiation time. The baseline utility (Definition 1) on the other hand, con-

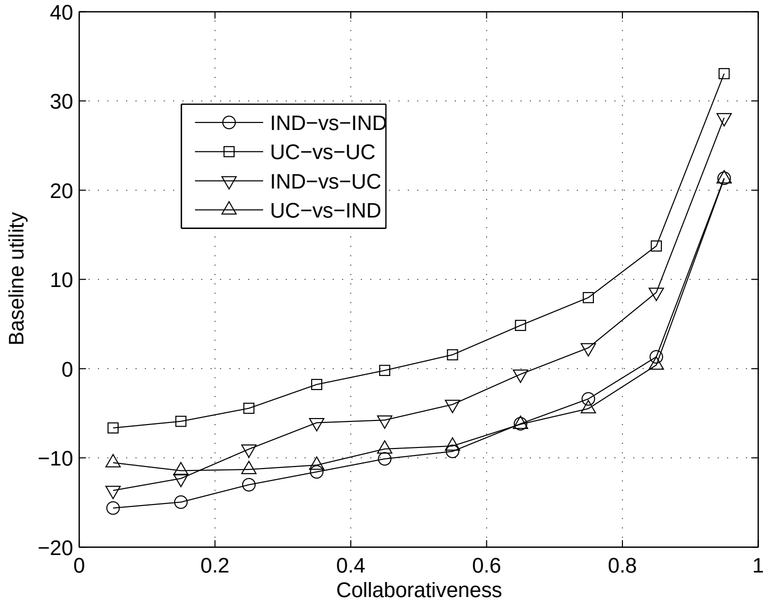


Figure 7.2: The baseline utility in the function of collaborativeness for the following strategy pairs: IND+A vs IND+A, IND+A vs UC+A, UC+A vs UC+A and UC+A vs IND+A.

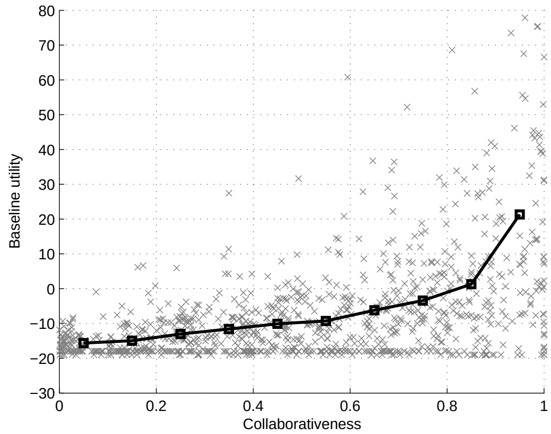
siders the time spent during negotiation as part of the cost. Certain strategies might choose to exit the difficult negotiation scenarios early even at the cost of an unrequited concession, which damages their relative utility, but can boost their baseline utility.

For the experiments describe in this subsection, we consider a negotiation round to take $t_r = 0.5$. The deadline for the IND strategy is $n_{max} = 40$, the conceding speed of the UC strategy is $\alpha = 0.05$. For the strategies which use collaborativeness analysis, we let the threshold $\Xi_{threshold} = 0.3$.

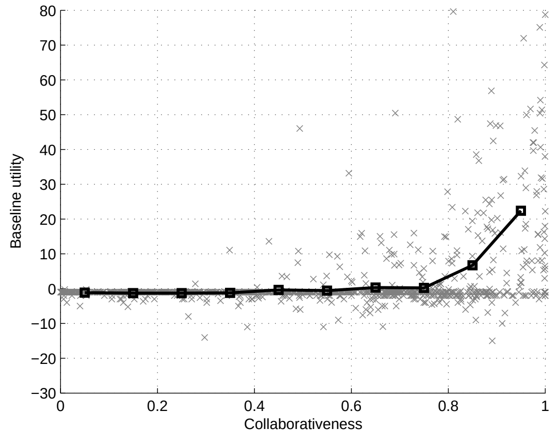
In the first set of experiments we compare the IND and UC strategies in all four possible pairings (IND+A vs IND+A, IND+A vs UC+A, UC+A vs IND+A and UC+A vs UC+A). We plot the the baseline utility in Figure 7.2. The first observation is that, same with the relative utility for all settings the baseline utility increases monotonically with the collaborativeness, but there is a significant variation among the negotiating strategy pairs. For the relative pragmatic utility the IND+A strategy always outperforms UC+A. For the baseline utility, however, the order is different, the best performance being obtained by the UC+A vs UC+A pairing. The baseline utility graphs shows how difficult is to obtain a positive negotiation result under the settings of our problem: the UC+A vs UC+A pairings yields negative average for $\Xi < 0.45$, but IND+A vs UC+A is negative for $\Xi < 0.65$ and UC+A vs IND+A and IND+A vs IND+A is negative for $\Xi < 0.85$! These negative values are a result of long negotiation sessions trying to obtain a better concession from the opponent, while loosing more on the time spent for each negotiation round.

Figure 7.3 shows the baseline utility for the IND+A, IND+A+CA, UC+A and UC+A+CA strategies when negotiating with opponents using the same strategy. In addition to the averages, these graphs also show the scatter plot of the individual negotiation results. The immediate observation is the significant improvement of the IND+A+CA and UC+A+CA strategies for the low collaborativeness values. While the scatter plot shows a large number of negotiations finishing in the negative for IND+A and UC+A, there are virtually none of them for IND+CA and UC+CA.

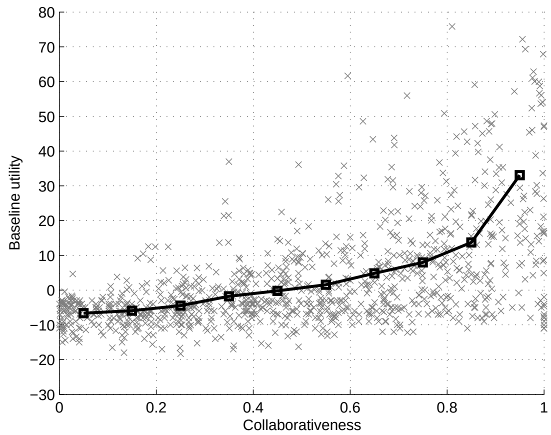
Another noteworthy feature is the visible concentration of points on the -20 horizontal line at Figure 7.3-a (IND+A vs IND+A). This line corresponds to the $t_r \cdot n_{max} = 0.5 \times 40 = 20$ value of the negotiations where the IND+A agent was forced to take the conflict deal after reaching the internal negotiation deadline.



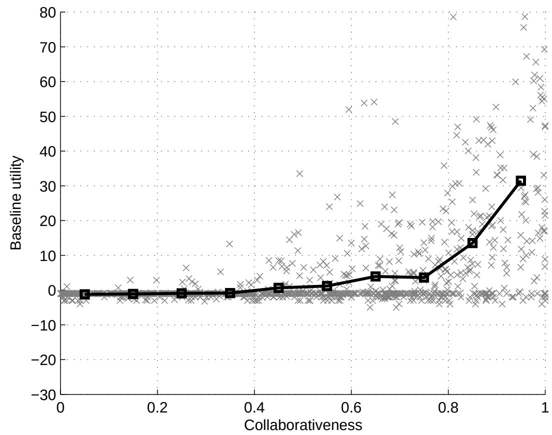
(a) IND+A vs IND+A



(b) IND+A+CA vs IND+A+CA



(c) UC+A vs UC+A



(d) UC+A+CA vs UC+A+CA

Figure 7.3: The baseline utility of the IND+A, IND+A+CA, UC+A and UC+A+CA agents negotiating with opponents using the same strategy.

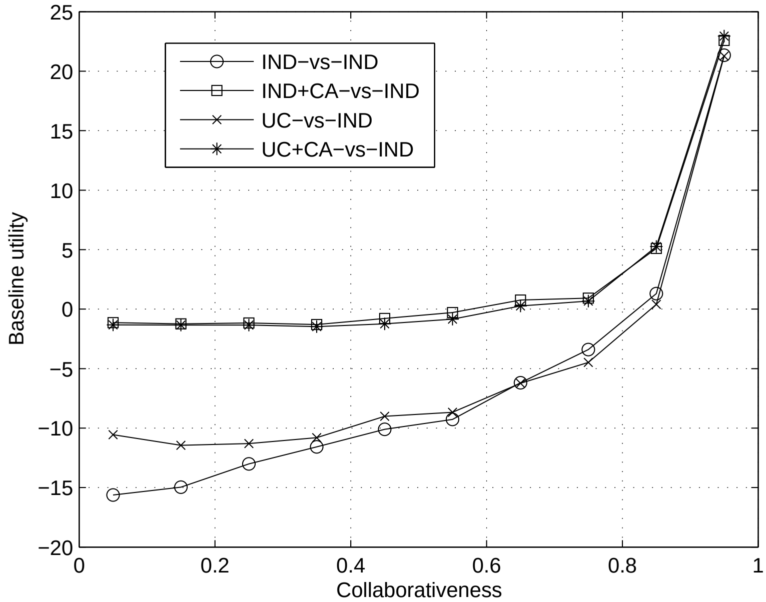


Figure 7.4: Baseline utility of the IND+A, IND+A+CA, UC+A and UC+A+CA strategies negotiating against an agent using the IND+A strategy.

A similar concentration of points can be found around the line corresponding to zero utility for the IND+A+CA and UC+A+CA graphs. These points correspond to the case when the collaborativeness analysis component dictated an early termination of the negotiation.

For a closer analysis of the relative performance, we ran a series of experiments where all the proposed strategies (IND+A, UC+A, IND+A+CA and UC+A+CA) negotiate against the same opponent, IND+A for Figure 7.4 and UC+A for Figure 7.5. The trend is similar for all the combinations on these graphs: the strategies augmented with collaborativeness analysis significantly outperform the other ones for low collaborativeness values, limiting their losses to the cost of the several negotiation rounds necessary to come up with an estimate. For scenarios of high collaborativeness, on the other hand, the performance is roughly equivalent.

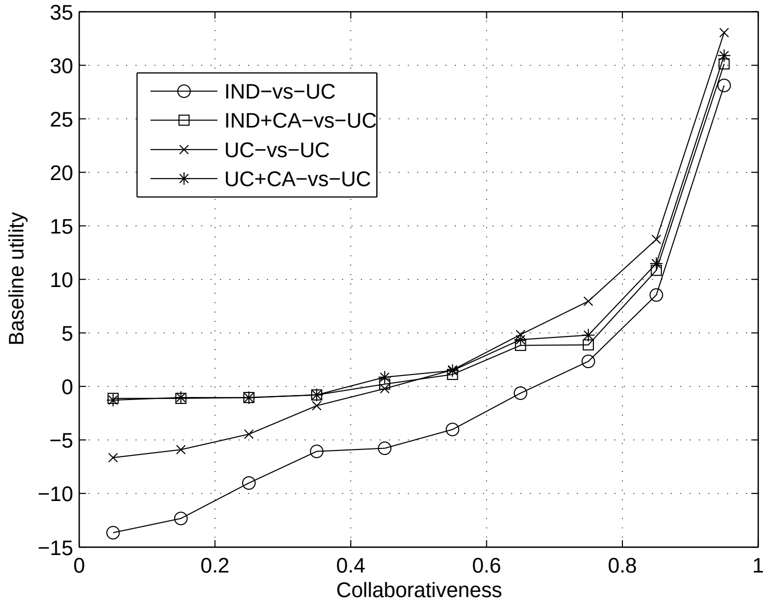


Figure 7.5: Baseline utility of the IND+A, IND+A+CA, UC+A and UC+A+CA strategies negotiating against an agent using the UC+A strategy.

In some cases, such as the UC+A vs UC+A and UC+A+CA vs UC+A in Figure 7.5, the CA version might perform slightly worse for the highest collaborativeness values. This phenomena appears because of the inaccuracies of the collaborativeness estimation.

7.2 Learning opponent through Bayesian learning

As we stated at the beginning of this dissertation, incomplete information is the default assumption. The self-interested negotiation partners disclose preferences only in the degree they believe that it allows them to reach a more favorable agreement. Naturally, a better knowledge of the opponent's preferences allows an agent to form better offers, and ultimately to reach a more favorable deal.

The preferences of the agent participating in spatio-temporal negotiation are defined in terms of physical properties such as current physical location, desired destination, current and maximum velocity, remaining fuel, desired trajectories and so on. This requires a different approach compared to worth oriented or task oriented domains.

In this section, we outline a technique which allows an agent to learn the preferences of the opponent agent. The negotiation protocol we assume is a simple exchange of binding offers (EBO) - that is, there are no arguments exchanged, the agent needs to infer the preferences of the opponent from its offers, or from the rejection of its own offers by the opponent.

Our approach is based on Bayesian learning which was previously used for multi-agent negotiations by Zeng and Sycara [ZS98], Li and Cao [LC04] and others. The agent updates its beliefs about the opponent's preferences after each negotiation round.

The main contributions of this section are the specific techniques which need to be used to calculate the posterior probabilities considering the spatial and temporal nature of the preferences, and the specific dependencies between the preferences. In addition, in contrast with most previous work in preference learning, we do not assume that the opponent uses a specific negotiation strategy.

The only assumptions about the opponent are those dictated by common sense: (a) that it does not make binding offers which are not feasible for itself (b) that is does not make binding offers which are not rational for itself (they are worse than the conflict deal) (c) that from a pool of available offers it presents the ones with the higher utility for itself before the ones with the lower utility, and (d) that it doesn't reject the offer which is better than the counter offer it plans to propose next round. Note that the third requirement does not necessarily imply a uniform concession. There is a very large space of possible strategies which verify these requirements. These four assumptions translate into three algorithms for the computation of the posterior probabilities in the Bayesian learning.

7.2.1 Bayesian learning during negotiation

The nature of the offer discloses some velocity information between agents. Specifically, for an offer by the opponent, the agent can easily calculate the common speed that the opponent wants to use to traverse the forest. This speed can not exceed the maximum speed of opponent, because it will not propose an offer which is not feasible for itself. In this way, to guess the speed of opponent, the learning agent just needs to calculate the maximum common speed from all the previous offers it received from the opponent. Moreover, it should add some time buffers in the splitting time field (if necessary) when proposing the next counter offer to the opponent.

To guess the source location and destination of the opponent, the map is divided into grid. The combination of a grid in the source area and another one in destination area is called a location model. The learning agent tries to guess the location model of the opponent, by updating the probabilities (belief) of all these combinations. Initially, each location model has equal probability, and the sum of these probabilities equals to one. From time to time, these probabilities are updated along the number of offers the learning agent receives from the opponent.

Bayesian learning is the classical method to update the belief based on evidences. Mathematically, the probability that the opponent is in the location model $\{sx, sy, dx, dy\}$ (the coordinates of the grid cells), when receiving a new evidence O_t (receiving an offer from opponent) can be calculated based on Bayes' theorem.

$$Pr(\{sx, sy, dx, dy\}|O_t) = \frac{Pr(O_t|\{sx, sy, dx, dy\})Pr(\{sx, sy, dx, dy\})}{\sum_{i,j,k,l=0}^{i,j,k,l=grid-1} Pr(O_t|\{i, j, k, l\})Pr(\{i, j, k, l\})} \tag{7.4}$$

where $grid$ is the number of pieces the learning agent divides the map in each dimension, and t is the order of the offers it receives from the opponent. The formula shows that the posterior probability of a location model can be calculated by the prior probability times the probability to propose the offer given the opponent is indeed in the specific location model, and then normalized by all the updated probabilities. The learning algorithm is shown in algorithm 4.

Algorithm 4 Algorithm for the learning agent

initialize all location models and assign them equal probability;
for $t = 1$ to $theEndOfNegotiation$ **do**
    get the opponent's offer O_t;
    for all location models $\{i, j, k, l\}$ **do**
        calculate $Pr(O_t|\{i, j, k, l\})$;
        updated posterior probability $Pr(\{i, j, k, l\}|O_t)$;
    end for
    normalize all the updated probabilities;
    propose next offer to opponent;
end for

7.2.2 Determine the posterior probabilities

In this subsection, we will discuss how the learning agent calculates $Pr(O_t|\{i,j,k,l\})$ - the probability to propose the offer O_t, given that the opponent is in location model $\{i,j,k,l\}$. First, we establish four basic rules according to the assumptions of opponent agent. We let the learning agent eliminate non-rational location models which break these rules. Next, the learning agent will calculate the expected utility of opponent at a specific negotiation round, and increase the probabilities of the location models whose actual utilities of the offer are close to the expected one. At last, we introduce a half Gaussian approach to overcome the case where the learning agent doesn't know the expected utility for the opponent.

The four basic rules

We are going to make four basic assumptions about the behavior of the opponent agent in the negotiation. First, the opponent will not propose an offer which is not feasible for itself. Second, the opponent will not propose an offer which is not rational for itself, (otherwise, it will arrive the destination later than its conflict deal). The third assumption is the opponent will propose a counter offer whose utility for itself is less or equal than the previous offers. This means that at each round of negotiation, the opponent should concede or at least insist on its last offer. The last assumption is that the opponent will accept the agent's offer if its utility is higher than the next counter offer. If the opponent in an assumed location model proposed an offer which breaks these rules, the learning agent will eliminate the possibility of that location model.

Practically, the value of $Pr(O_t|\{i,j,k,l\}) = 0$ if the opponent was assumed at location model $\{i,j,k,l\}$ but its last O_t breaks the four basic rules. All the other location models in the learning agent's belief share the same probability. Next, the learning agent will continue to discriminate these rational models and finds the one more likely.

Updating belief based on expected utility

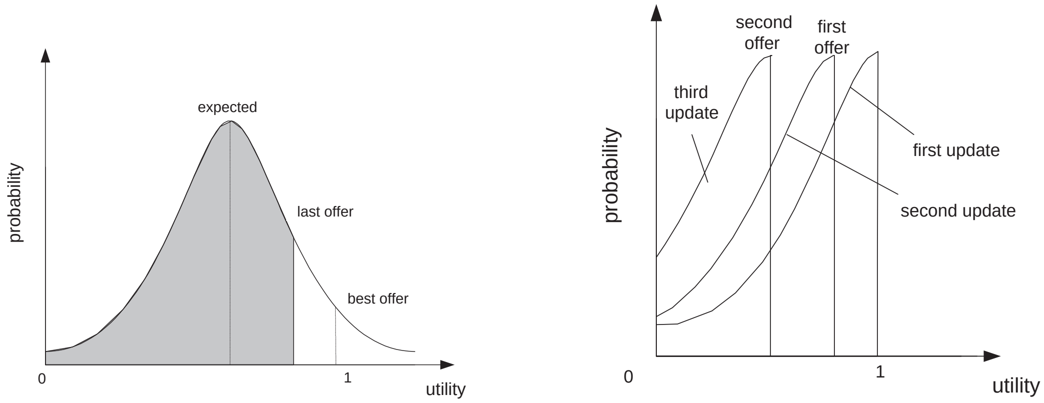
A self-interested agent will not only act rational, but also propose the most profitable offers at first, and concede to less profitable ones later. Using this idea, the learning agent can calculate the expected utility at a specific negotiation round, and assign more probabilities to those location models for which the utility of the offer is close to the expected one. In practice, the learning agent assumes that the opponent proposes offers with utilities starting from 1.0 at the first call and linearly decreasing during the negotiation.

$$EU(t) = 1 - \alpha \times t \tag{7.5}$$

where t is the order of the offers by the opponent and α is the conceding speed. At each negotiation round, the location model whose utility of the offer O_t is close to $EU(t)$, will have its probabilities increased based on the Gaussian p.d.f.

$$Pr(O_t|\{i,j,k,l\}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(U_t(O_t,\{i,j,k,l\})-EU(t))^2}{2\sigma^2}} \tag{7.6}$$

where $U_t(O_t, \{i,j,k,l\}))$ is the utility of the opponent's offer O_t when it is assumed in location model $\{i,j,k,l\}$, and σ is the coefficient of confidence. There are several approximations for



(a) Expected utility method

(b) Half Gaussian method

Figure 7.6: Two methods to discriminate location models in the learning agent: 7.6(a): it updates belief based on Gaussian p.d.f which center at the expected utility and 7.6(b): it updates the probabilities based on half Gaussian p.d.f with the center at the utility of last offer.

this approach. The first one is we transfer a four-dimensional vector (offer O_t) into a value (utility U_t) and assume they have the same posterior probabilities.

$$
\begin{aligned}
& Pr(O_t|\{i,j,k,l\}) \\
& = \tfrac{Pr(U_t|\{i,j,k,l\}) \times Pr(O_t|U_t,\{i,j,k,l\})}{Pr(U_t|O_t,\{i,j,k,l\})} && (Bayes' theorem) \\
& = Pr(U_t|\{i,j,k,l\}) \times Pr(O_t|U_t,\{i,j,k,l\}) && (definition\ of\ utility) \\
& = Pr(U_t|\{i,j,k,l\}) && (assumption)
\end{aligned}
$$

The equation assumes that $Pr(O_t|U_t,\{i,j,k,l\}) = 1$. In general, an agent may find many offers given a specific utility, and the assumption is not true for those strategies which want to try out every possible offer before conceding the utility. However, considering the negotiation time is crucial, we assume the opponent can only select one offer given a specific utility.

Another approximation for this approach comes from the four basic rules. The learning agent eliminates the probability of non-rational location models whose utilities are negative or greater than the utility of the opponent's last offer. Such elimination cuts off the Gaussian p.d.f (see Figure 7.6(a)), and the integral of the remaining part doesn't equal one. The assumption here is we ignore these parts because all the probabilities will be normalized later, and we just need a discriminant value to judge the distance between the actual utility and the expected one. In the mean time, we can also change the variance of Gaussian p.d.f to reduce the impact of this approximation.

The main deficiency of this approach is the difficulty to find a correct conceding speed to calculate the expected utility. If the opponent uses a different strategy which is not linear concession in utility, the learning agent may make a wrong guess. To overcome this problem, we need to model the opponent's strategy and calculate the expected utility based on the probabilities of strategy models [FP08] (we leave it in the future work), or we can apply it in a save way which we will discuss next.

Updating belief based on the half-Gaussian distribution

The idea of this approach is that an agent will concede step by step. At each step, it will give up a small amount of utility and see if the opponent accepts it. In this way, if the opponent which is assumed in a location model proposes two adjacent offers which have a big difference in utilities, the probability that the opponent is in that location model should be small.

Figure 7.6(b) depicts the way the learning agent calculates the conditional probability $Pr(O_t|\{i,j,k,l\})$. As we discussed above, the offer O_t is first transferred into utility U_t, given the assumption that the opponent is in location model $\{i,j,k,l\}$. Then, the learning agent calculates the probability of the offer based on utility and half Gaussian p.d.f, in which the mean of the Gaussian is at the utility of the last offer given the opponent is in the same location model.

7.2.3 Experimental study on applying Bayesian learning in the CRF game

7.2.3.1 The performance of learning

In this experiment, we focus on the accuracy of learning by comparing the opponent's actual location model with its probability in the learning agent's belief. At first, we test a typical scenario where the opponent is fixed in the center of a pair of specified grids (see Figure 7.7). The opponent is MCS agent with parameter of (2,2). The learning agent is located at the lower corners of the forest, and use different methods to update the posterior probabilities.

Figure 7.8 shows the updating progress in the learning agent's belief. First of all, at the beginning, all 81 location models are initialized as equal probabilities. When the learning agent uses four basic rules to update the posterior probabilities, some of location models are gradually eliminated. At the end of the learning, there are still 9 models in agent's belief, so they share equal probabilities (see Figure 7.8(a)). When we set up a conceding speed and ask

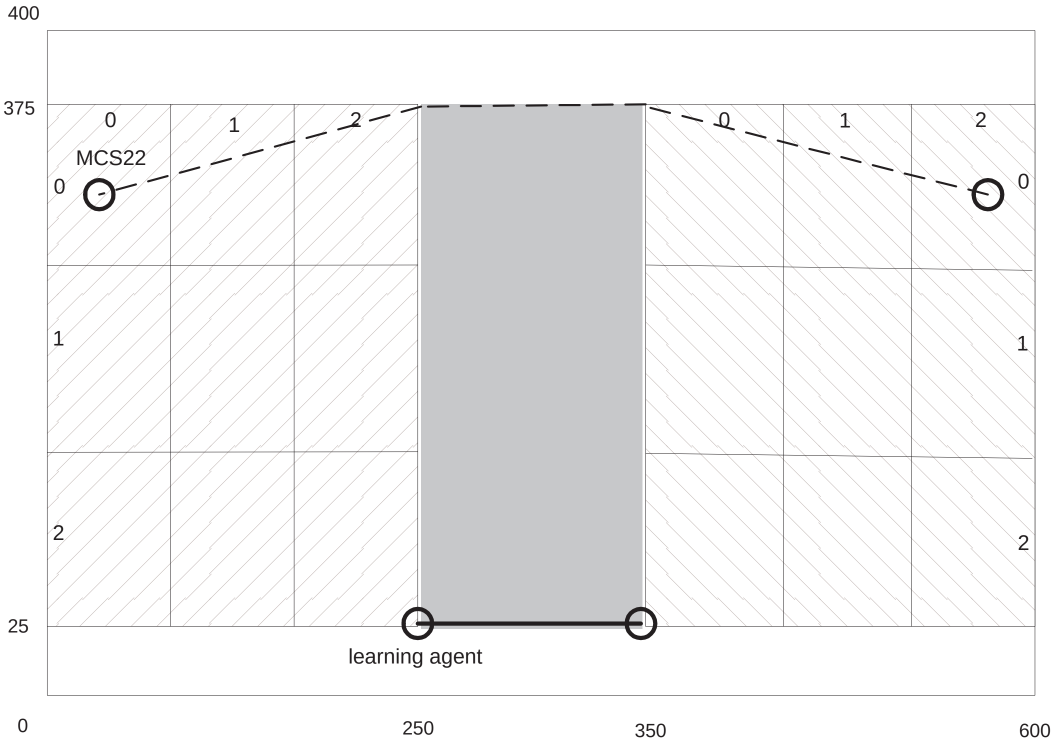
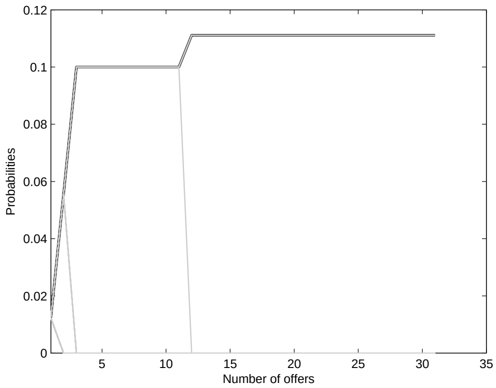
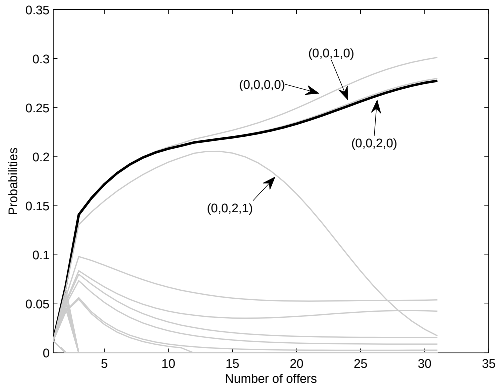


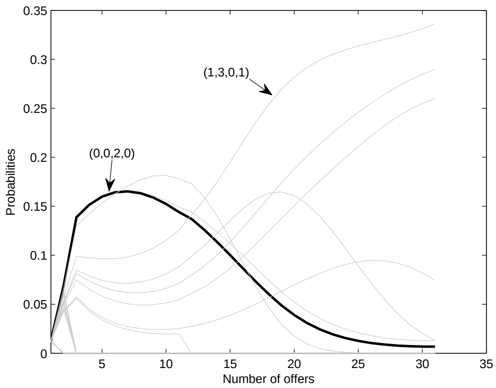
Figure 7.7: A typical scenario: both the source and the destination area is divided into 3 × 3 girds, which corresponds to 81 location models. The opponent agent is located at the center of grid (0,0) and wants to move to the center of grid (0,2) with the speed of 1.0. The learning agent is located at the lower-left corner of the forest, insists its best offer until the end of negotiation.



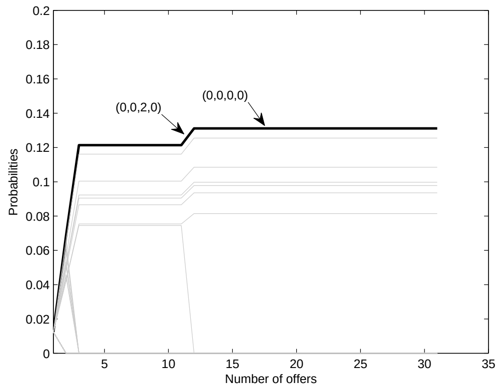
(a) The four basic rules



(b) Expected utility (correct)



(c) Expected utility (incorrect)



(d) Conceding pace method

Figure 7.8: The probabilities are updated along the number of offers from the opponent, the bold line is the opponent's actual location model in the learning agent's belief. The learning agent use: 7.8(a): four basic rules, 7.8(b): expected utility with correct conceding speed, 7.8(c): expected utility with incorrect conceding speed, 7.8(d): half Gaussian method, to determine the posterior probabilities.

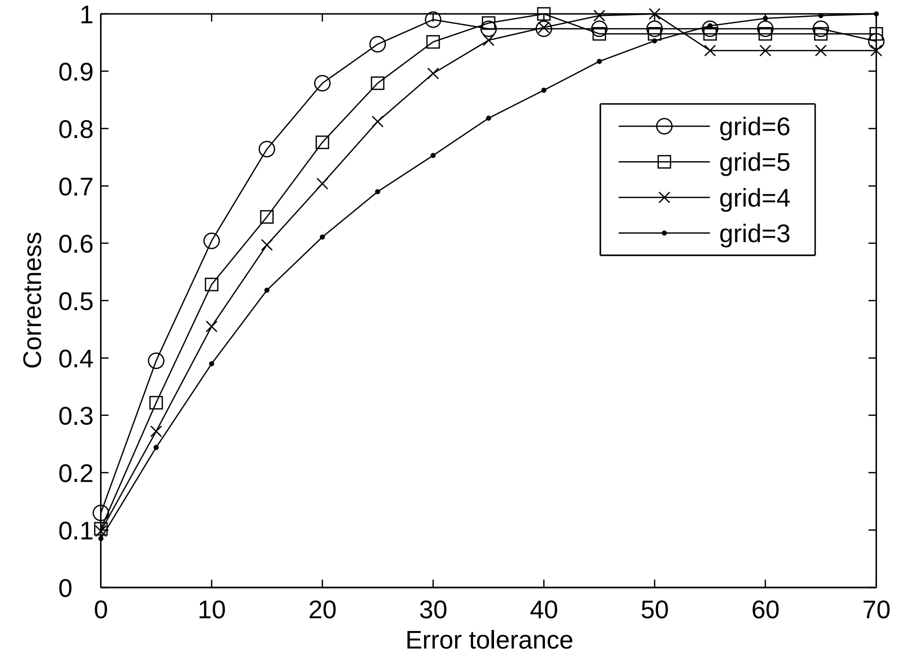
	MCS C_{meet}=2, C_{split}=2	UC α=0.02
grid=3, 81(models)	5.271	4.099
grid=4, 256(models)	9.731	7.016
grid=5, 625(models)	17.733	11.9936

Table 7.1: The number of location models remained in the belief of the learning agent.

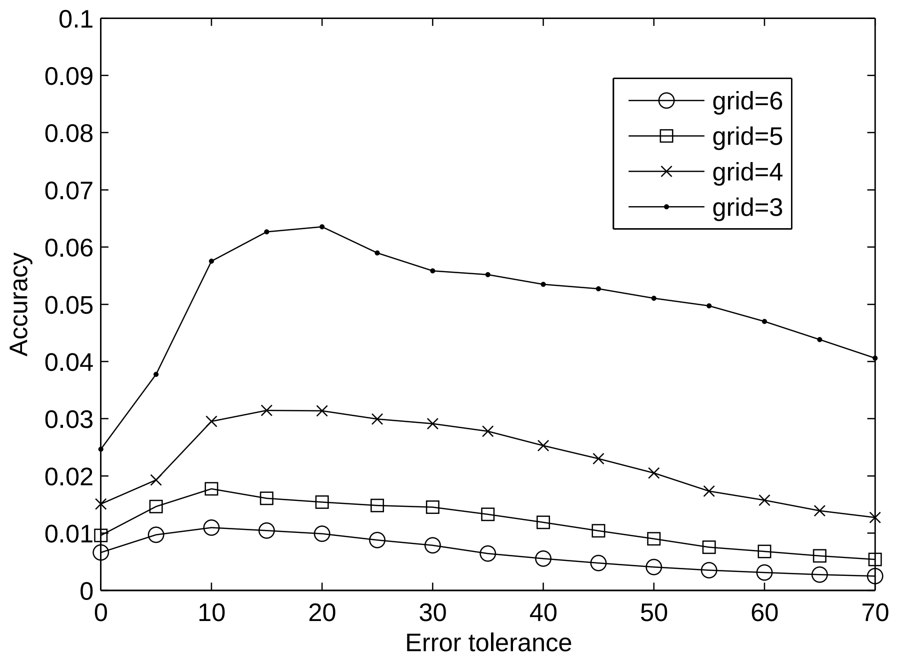
the learning agent to update probabilities based on Guassian distribution. The 9 remaining location models are further discriminated (see Figure 7.8(b)). Unfortunately, the deficiency of using expected utility is disclosed when we assign an incorrect conceding speed when the learning agent calculates expected utility at each negotiation step (see Figure 7.8(c)). At last, half Gaussian method gives a relatively compromised outcome, with the probability of the correct model ranging in the middle of the previous cases (see Figure 7.8(d)).

In the next experiment, we enumerate all possible combinations where the opponent's source and destination are at the centers of grids. We let the opponent use the MCS and UC strategy and the learning agent use the four basic rules. We calculate the averaged number of opponent models remained in the belief over all the combinations of grids and we increase the grid resolution to see the trend (see Table 7.1).

The next question is how to decide the error tolerance if the opponent is not at the centers of the grids. Possibly, the learning agent might eliminate the correct opponent model if the opponent is not at the center of the grids. This is also a trade-off between the correctness and the accuracy of learning. The correctness is if the correct model is still in the agent's belief. The accuracy is the probability of that correct model after the learning. Intuitively, if the error tolerance is too small, the correct location model may be eliminated, so its probability



(a) correctness



(b) accuracy

Figure 7.9: The values of correctness and accuracy in the function of error tolerance in 1000 random generated scenarios.

will be zero. On the other hand, if it is too large, the number of location models remained in the belief is also large, so the probability of the correct location model will be small too. In the next experiment, we generate 1000 random scenarios, we let the learning agent use the four basic rules negotiating with a MCS opponent. We calculate the correctness of these 1000 learnings, and their averaged accuracy. We change the value of error tolerance as well as the grid resolution in the function of correctness and accuracy (see Figure 7.9). The figure shows, by increasing the error tolerance in a specific value, both correctness and accuracy are balanced.

With a balanced error tolerance, we plot the performance of learning in these 1000 scenario when the agent use different approaches to update the posterior probabilities (see Figure 7.10).

7.2.3.2 Performance of negotiation with or without learning

To be better evaluate the performance of learning in negotiation, we firstly design a full knowledge agent: Uniform Concession in Full knowledge (UCF) as the benchmark strategy. The UCF strategy is similar to UC when proposing offers but it works as a supervisor (it knows the whole scenario in advance). When proposing the next offer, the UCF agent will choose the offer in the current subpool which provides the opponent best utility. If the opponent doesn't accept the offer, it will decrease to the next subpool until there is no

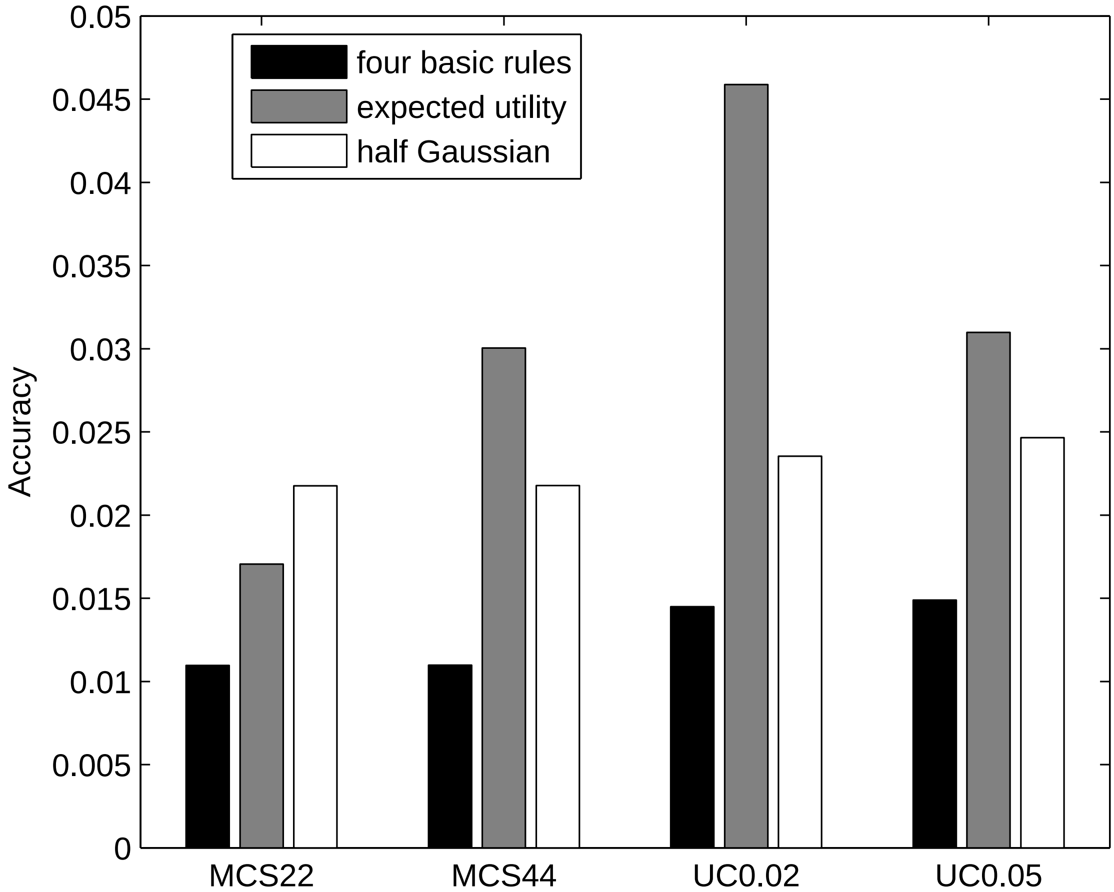
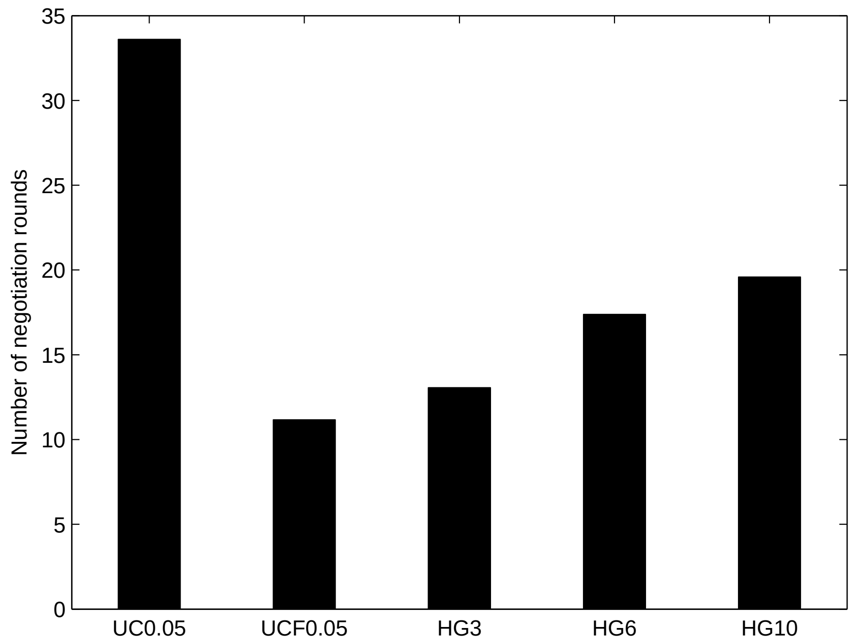


Figure 7.10: The statistical study about the three learning approaches, when the opponent use the MCS22, MCS44, UC0.02, and UC0.05 strategies respectively. The results come from 1000 random generated scenario with learning agent's grid number of 6, error tolerance of 10. The learning agent which uses expected utility method assume the opponent's conceding speed is 0.03.

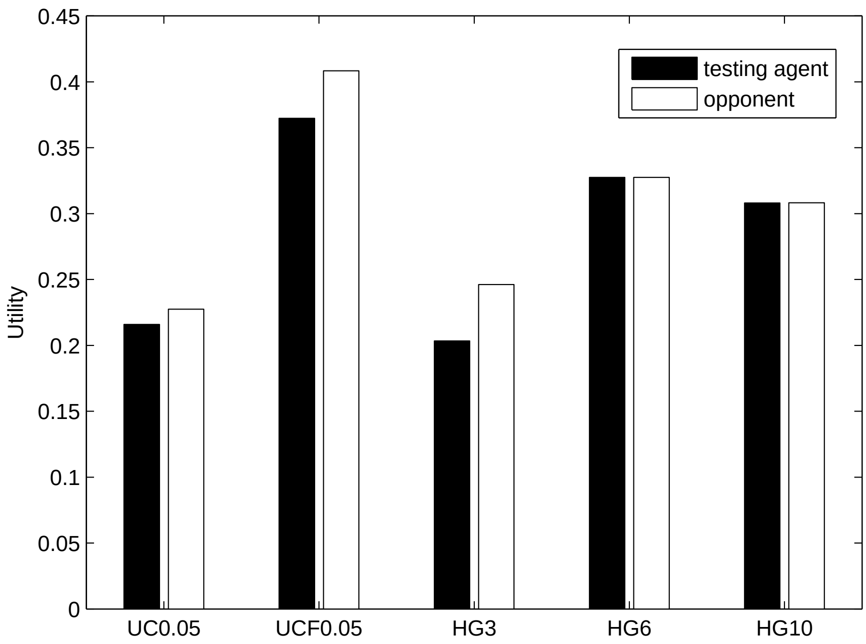
subpool exist in the next round. The agent will quit the negotiation if there is no agreement at that time.

To apply the output of learning in the negotiation, we design the Uniform Concession with Learning (UCL) agent. The UCL agent calculate its next offer pool in the same way with the UC agent (according to the utilities). The difference is it will guess the opponent as if it were in the most probable location model of its current belief and provide it the best utility as UCF does. If the opponent doesn't accept it, it will not concede to the next subpool but continue to update beleif. After a certain round of learning (a certain amount of probabilities in belief have been searched), it decreases to the next subpool and does the same thing again. The story behind this approach is that the learning agent initially doubts the correctness of its own belief before it concedes the utility level. In the other word, it updates the subjective belief before concede the objective scenario it can not control (such as the opponent's strategy, and the opponent's location).

In Figure 7.11, we compare the averaged number of negotiation rounds, and the averaged relative utility the agent gains, for the UCF agent (an agent with full knowledge), the UC agent (an agent without knowledge and without learning), and the UCL agent (an agent which uses half Gaussian to update belief, without knowledge but with learning). Each of them negotiates with the opponent UC agent. From the experiment, we can see: (a) the number of negotiation rounds is dramatically decreased if the agent is learning; The more location models in its initial belief, the more number of negotiation round it takes. (b) the



(a) Number of negotiation rounds



(b) Averaged utilities

Figure 7.11: The benefit of learning: UC0.05, UCF0,05, UCL3, UCL6 and UCL9 negotiate with another UC0.05 in 1000 random scenarios

utility of the deal is also improved if the agent learns; and (c) the relative utility is increasing when the learning agent increase the grid resolution of the map.

CHAPTER 8
ACTING WHILE NEGOTIATING

The assumption so far is that the agents are still immobile during the negotiation. In the CRF game where negotiation happens in the physical time, the more rational agents would move while negotiating. The work described in this chapter represents a step towards bringing convoy negotiation closer to a more realistic setting. Rather than assuming that the agents are negotiating instantaneously, we assume that the negotiation process is happening in physical time, during which the agents can take real world actions, such as moving towards their destination, their expected meeting point or other locations.

Allowing agents to move while negotiating asks for a pair of action and negotiation strategies in the agent. It also generates the problem that the opponent is not in the original source location in the learning agent's belief. The relationship between the action strategy and negotiation strategy is complex. A good action strategy will consider the current status of negotiation; in its turn, the actions taken by the agent will change the value of the exchanged offers.

In this chapter, we introduce the strategy which apply particle filter in a time evolving model. Through a series of experiments, we study the interaction between the negotiation and action strategies and compare the performance of the proposed strategy pairs in incomplete information scenarios.

8.1 The selfishness-optimism meta-strategy

To capture the relationship between the action strategy and the negotiation strategy into an easy-to-understand framework, we propose a technique which integrates the offer acceptance decision and the action strategy into a single meta-strategy. This *selfishness-optimism meta-strategy* (see Algorithm 5) does not define the offer formation mechanism; this needs to be provided separately, and is normally inherited from non-AWN strategies.

Algorithm 5 The selfishness-optimism meta-strategy

receive(O^B_{i-1});
$Belief(t) \leftarrow Belief_{update}(Belief(t-1), O^B_{i-1})$;
if $isFeasible\left(O^B_{i-1}\right)$ and $U(O^B_{i-1}) \geq \lambda$ **then**
    send$\left(O^B_{i-1}\right)$; // form agreement
else
    $O^A_i = NextOffer\left(Belief(t), \lambda\right)$;
    if not $isFeasible\left(O^A_i\right)$ **then**
        send($\emptyset$); // conflict deal
    else
        send(O^A_i);
    end if
end if
$L^A(t)$ = move($L^A(t)$, Belief(t), γ);

The *selfishness* λ is the lowest utility of the offer, as defined by Equation 7, which the agent is ready to accept. A fully selfish agent ($\lambda = 1$) will only accept its ideal offer, a fully benevolent agent ($\lambda = 0$) will accept any rational offer.

The *optimism* γ governs the agent's movement and represents the amount of hedging between moving towards its own latest offer versus the conflict deal location. A fully pessimistic agent ($\gamma = 0$) assumes that there will be no deal and move on the conflict deal trajectory.

The reader might notice that this meta strategy can be immediately generalized by making the λ and γ parameters variable over the course of the negotiation. An agent, seeing that the opponent conceded too readily, might decide to drive a hard bargain by increasing its selfishness. An agent might make its optimism dependent on an external machine learning system which predicts the likelihood of a deal. A particularly Machiavellian agent might even make offers only to confuse the opponent and move to a predicted deal location which is far from its current offer. For the remainder of this section, we will assume agents with the λ and γ parameters fixed and determined at the beginning of the negotiation.

8.2 Minimal opponent models

We assume a zero-knowledge environment: the only information the agents have about each other is extracted from the offers. Under these conditions, one of the main challenges of any convoy negotiation is offer formation. With four issues in every offer, the negotiation space is very large. There is no natural ordering of the offers, thus an agent can never be sure whether the offer it made is a concession to the opponent or not, and whether it is feasible or rational for the opponent. Finally, as the opponent is also acting while negotiating, the utility, rationality and feasibility of the offers changes with the passing of time and the actions of the opponent. All this makes it important that the agent as part of its negotiation strategy build a model of the opponent.

In the following we consider some initial information which can be extracted from the initial offers of the agents. An offer does not immediately identify the agent's source and destination, even if the agent offers its own ideal trajectory. As the same in section 7.2, the factor which is relatively easy to identify is the speed capability of the agent. As every offer is binding, the first offer made by an agent will identify a minimum value on the agent's speed capability based on the speed on the common trajectory portion. Unless the agent is engaged in deceptive practices, this first offer will be based on its maximum possible speed.

The agent making the second offer can find itself in one of two possible situations. It can find that the opponent's speed is larger than its own. Then it needs to structure its counter-offer based on its own, lower speed. On the other hand, if it finds the opponent's speed to be smaller than its own capability, it will make an offer assuming the opponent's speed for the common part of the trajectory, without disclosing its own higher capabilities. Either way, by the end of the first offer exchange, the agents will know their maximum common speed, and will use this in all subsequent offers. Thus, the remainder of the offers will always be feasible for the common portion of the trajectory. It is, however, much harder to determine the current location of the opponent agent.

8.3 Opponent modeling with particle filters

As the opponent is moving while negotiating, in our case, opponent modeling requires not only learning the initial parameters, but also maintaining a dynamically evolving model of the opponent, a problem of *probabilistic reasoning over time*. In this section we describe the PF strategy which uses a Sampling-Importance-Resampling (SIR) particle filter to update its beliefs about the opponent, then uses a K-Means clustering technique to extract a likely hypothesis on which the offer formation is based.

The PF strategy represents its knowledge about the opponent as a cloud of weighted particles. In the following we discuss (1) the particle representation, (2) the prediction model, describing how the particles evolve in time and (3) the sensor model, which describes how observations (which in our case are offers made by the opponent) affect the weight of the particle.

The particle representation

A particle should contain all the information the learning agent needs to know about the opponent. We represent the particle $\mathbf{X}_t$ at time t as a vector of its opponent's current state:

$$\mathbf{X}_t = \langle L_{src}, L_{crt}, L_{dest}, S_{id} \rangle$$

where L_{src} is the source location, L_{crt} is the current location, L_{dest} is the destination, and S_{id} is an identifier of the strategy used by the opponent. The strategy is chosen from a set of discrete strategies.

The prediction model

At every negotiation round, the particle $\mathbf{X}_t$ is updated from its previous state $\mathbf{X}_{t-1}$ using the following equations:

$$\mathbf{X}_t = \begin{cases} L_{src}(t) = L_{src}(t-1) + \xi_{src} \\ L_{dest}(t) = L_{dest}(t-1) + \xi_{dest} \\ L_{crt}(t) = f(S_{id}, L_{crt}(t-1)) + \xi_{current} \\ S_{id}(t) = S_{id}(t-1) \end{cases}$$

where $f(.)$ is a function to calculate the next location according to the opponent's strategy S_{id} and its former location $L_{crt}(t-1)$ and $\xi_{.}$ is random noise generated from the two-dimensional normal distribution accounting for the uncertainty of the estimation.

The sensor model

The particle weights are updated with every new observation. For each particle, the PF agent calculates the probability $Pr(\mathbf{O}_t|\mathbf{X}_t^i)$ that a hypothetical opponent described by the particle would make the specified offer. To do this, we first calculate the offer which would have been made by the agent described by the particle $O_{exp}(X_t^i)$ and then calculate the probability based on the difference of the real offer from the expected offer:

$$
\begin{aligned}
Pr(\mathbf{O}_t|\mathbf{X}_t^i) &= Pr(O_t|O_{exp}(X_t^i)) \\
&= g_4(y_m, t_m, y_s, t_s|y_m^{exp}, t_m^{exp}, y_s^{exp}, t_s^{exp}) \\
&= g(y_m|y_m^{exp})g(t_m|t_m^{exp})g(y_s|y_s^{exp})g(t_s|t_s^{exp})
\end{aligned}
$$

In the formula, $(y_{meet}, t_{meet}, y_{split}, t_{split})$ is the actual values in opponent's last offer $\mathbf{O}_t$. $g_4(.)$ is the four-dimensional Gaussian p.d.f which centers at expected offer $O_{exp}(X_t^i)$ and with specific coefficient matrix.

$$
w_i(t) = Pr(\mathbf{O}_t|\mathbf{X}_t^i)w_i(t-1)
$$

The particle weights are normalized after the update, and if the estimate of effective number of particles

$$
\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{P}(w_i)^2}
$$

is less than the threshold $N_{threshold}$, we resample using the stratified resampling algorithm.

8.4 Offer formation for the particle filter

Algorithm 6 describes the calculation of the next offer by the PF agent. First, we associate an offer to every particle. By making the assumption that the particle is correct, we generate the offer the same way as if we would have a full-knowledge negotiation: the offer will be feasible to both agents and have a utility larger than their respective selfishness levels. If more than one such offer can be generated, we choose the one which is closest to the opponent's last offer.

Algorithm 6 Offer formation for the Partical Filter strategy

for all particles i **do**
    search all O^i where $U^{(A)}(O^i) \geq \lambda$ and $U^{(B)}(O^i) \geq \lambda_i$;
    if no any O^i **then**
        $O^i_{best} \leftarrow null$;
    else
        $O^i_{best} \leftarrow \arg\max U_{opponent}(O^i)$;
    end if
end for
if no particle has O^i_{best} or $\sum w_i \leq threshold$ **then**
    return $O_{next} \leftarrow null$;
else
    cluster all particles whose $O_{best} \neq null$;
    calculate weights of all clusters;
    find the most weighted cluster j;
    return $O_{next} \leftarrow O_{ave}(j)$;
end if

If none of the particles has a feasible associated offer, the PF agent breaks the negotiation. Otherwise the agent proceeds to choose an offer based on the offers associated with the particles. Calculating the mean across all the particles is not a good choice, as the particles might represent disjoint hypotheses. By taking the average over the complete set of particles, the resulting estimate might fall in the low probability zone between hypotheses.

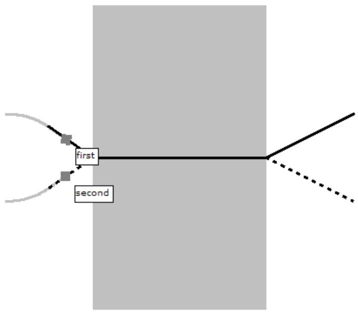
Our approach is to perform K-Means clustering on all the particles which have assigned offers. The distance metric used is the sum of squared difference between the issues. The cluster with the highest sum of weights is selected for offer formation. The averaged offer of the selected cluster will become the next counter-offer to the opponent.

8.5 Experimental study on applying the particle filter in the CRF-AWN problem

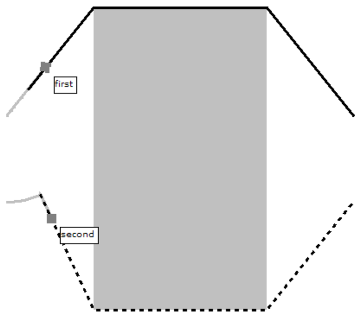
8.5.1 The influence of the selfishness and optimism on the agent trajectories

To understand the impact of the selfishness and optimism settings on the behavior of agents, we have run a series of experiments. We considered a scenario where a mutually advantageous deal is possible. The size of the map is 600×400, with the forest located at (200,25) with the size of 200 ×350. Agent A moves from (100,150) to (500,150) with the speed of 1.0, agent B with the fixed values of $\lambda = 0.6$ and $\gamma = 1$ moves from (100,250) to (500,250) with the speed of 1.0. Both agents use the MCS strategy($C_m = 2, C_s = 2$) to calculate the next offer. This is a "hard" scenario, because the social deal is only marginally better than the conflict deal.

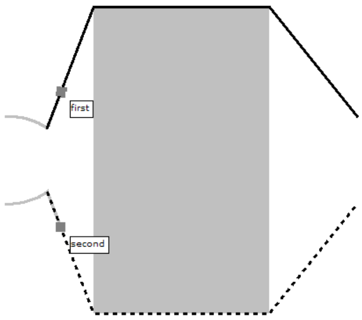
Figure 8.1 shows the path of the agents for four different settings of the selfishness and optimism for agent A. As the MCS strategy does not depend on the current location, the actual offers exchanged are identical. Interestingly, however, in cases (a) and (d) the agents agreed to form a convoy, while for (b) and (c) they did not. Figure 8.1-a shows agent A with $\lambda_A = 0.6$ and $\gamma_A = 1$, that is, of average selfishness but fully optimistic. The agent moves towards its own offer at every step which results in a curving trajectory as the offer evolves. As the agents are getting closer and closer together, the utility of their respective offers keeps increasing, thus a deal is eventually reached.



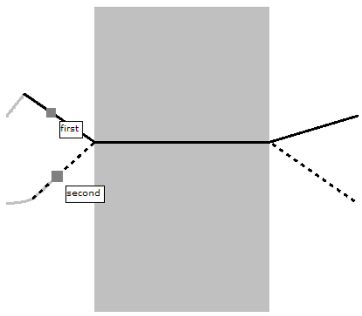
(a) Fully optimistic



(b) Fully pessimistic



(c) Very selfish



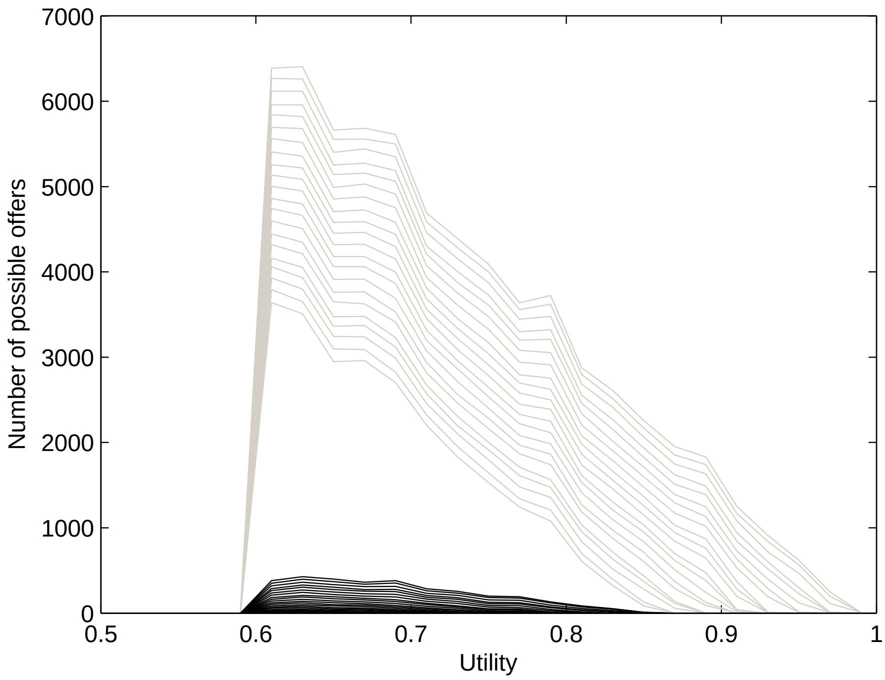
(d) Very generous

Figure 8.1: The influence of the selfishness and optimism to the course and the outcome of the negotiation. The meta-strategy of agent B is fixed to $\lambda_B = 0.6$ and $\gamma_B = 1$. The values for agent A are: (a) $\lambda_A = 0.6$, $\gamma_A = 1$, (b) $\lambda_A = 0.6$, $\gamma_A = 0$, (c) $\lambda = 0.8$, $\gamma = 1$ and (d) $\lambda = 0.2$, $\gamma = 1$.

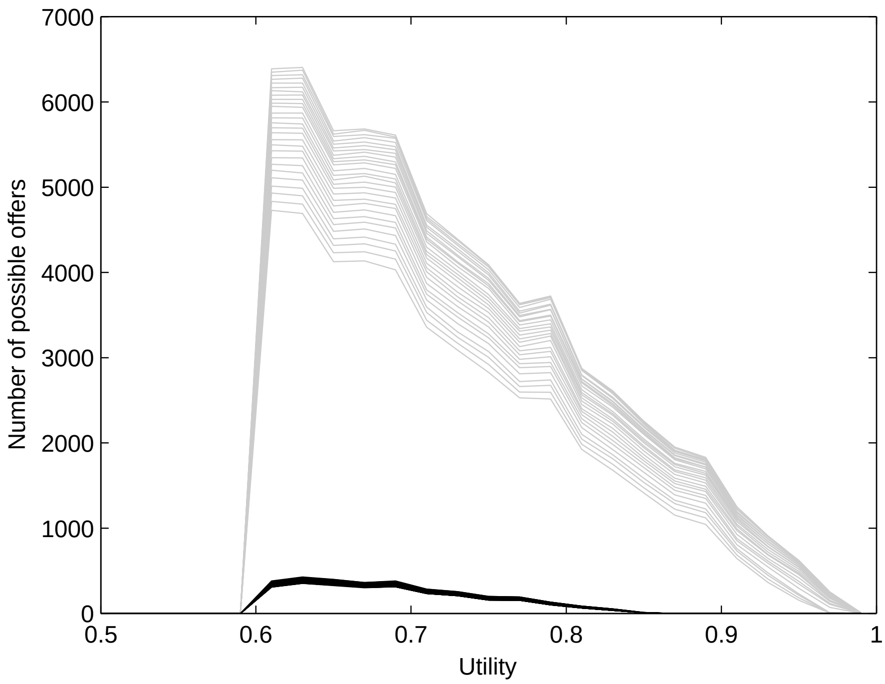
In Figure 8.1-b agent A is fully pessimistic and of average selfishness ($\lambda_A = 0.6$, $\gamma_A = 0$). Agent A moves in a straight line towards the conflict deal, making both its own and the opponent's offers less and less valuable, despite the concessions of the opponent. Finally, the offer which the agent needs to make according to its strategy becomes of lower utility than its selfishness, the negotiation is terminated, and the opponents move on the conflict deal trajectory. Note that agent B actually ended up on a trajectory which is worse than the original conflict deal.

Figure 8.1-c shows a run with A being fully optimistic but of high selfishness ($\lambda_A = 0.8$, $\gamma_A = 1$). The trajectories are initially similar to case (a), however, through a series of concessions, agent A will reach a point where its next offer will have an utility smaller than its selfishness. At this point A breaks of the negotiation and moves to the conflict deal. In this case *both* agents end up on trajectories which are worse than the original conflict deal.

Finally, Figure 8.1-d shows a case when A is fully pessimistic but of low selfishness ($\lambda_A = 0.2$, $\gamma_A = 0$). Despite the fact that it starts to move towards the direction of the conflict deal, A and B successfully form a deal because A will accept a relatively low utility rational offer. Thus A will reverse its course and move towards the collaborative deal. Note that A had lost some utility by making the "detour" towards the conflict deal.



(a) Fully pessimistic



(b) Fully optimistic

Figure 8.2: Evolution of the histograms of offer pool (gray lines) and the supervisor's pool (black lines) function of the utility. (a) $\gamma = 0$ (fully pessimistic) and (b) $\gamma = 1$ (fully optimistic). For both cases, the $\lambda = 0.6$.

8.5.2 The influence of the action strategy on the offer pool

Let us consider a negotiation turn where agent A needs to make an offer. We call the agent A's *offer pool*, the set of offers which are rational and feasible for A. The *supervisor's pool* is the set of offers which are feasible and rational for both A and B. Some strategies, such as UC generate the agent's pool explicitly. The supervisor's pool can not be computed by the agents in partial knowledge negotiations.

In the acting while negotiating problem, both the agent pool and the supervisor's pool decreases at every negotiation round, as some offers become unfeasible, as a result of the passage of time and the action strategy of the agent.

One way to characterize the agent and the supervisor pools is to consider the histogram of the offers in function of their utility. Figure 8.2 plots the evolution of these histograms over the negotiation scenario described in the previous section. Series of gray lines show the agent's offer pool, and black lines the supervisor's pool. Figure 8.2-a considers a fully pessimistic agent. As expected, the agent offer pool shrinks at every iteration. Furthermore, the maximum utility of the agent's offer pool also becomes lower at every iteration, reflecting the fact that by moving on the conflict deal trajectory, the agent is reducing its own choices. The supervisor's pool is shrinking on its own as well, and eventually becomes empty.

Figure 8.2-b considers a fully optimistic agent. We note that the offer pool is still shrinking at every iteration, but the amount of decrease is smaller. Furthermore, the maximum possible utility remains very close to 1.0 during the negotiation, because the agent optimisti-

cally moves towards these high utility offers. We also notice that the rate of shrinking of the supervisor's pool is much slower than in the pessimistic case. Most of the offers which were feasible at the beginning of the negotiation remain feasible if the agent acts optimistically.

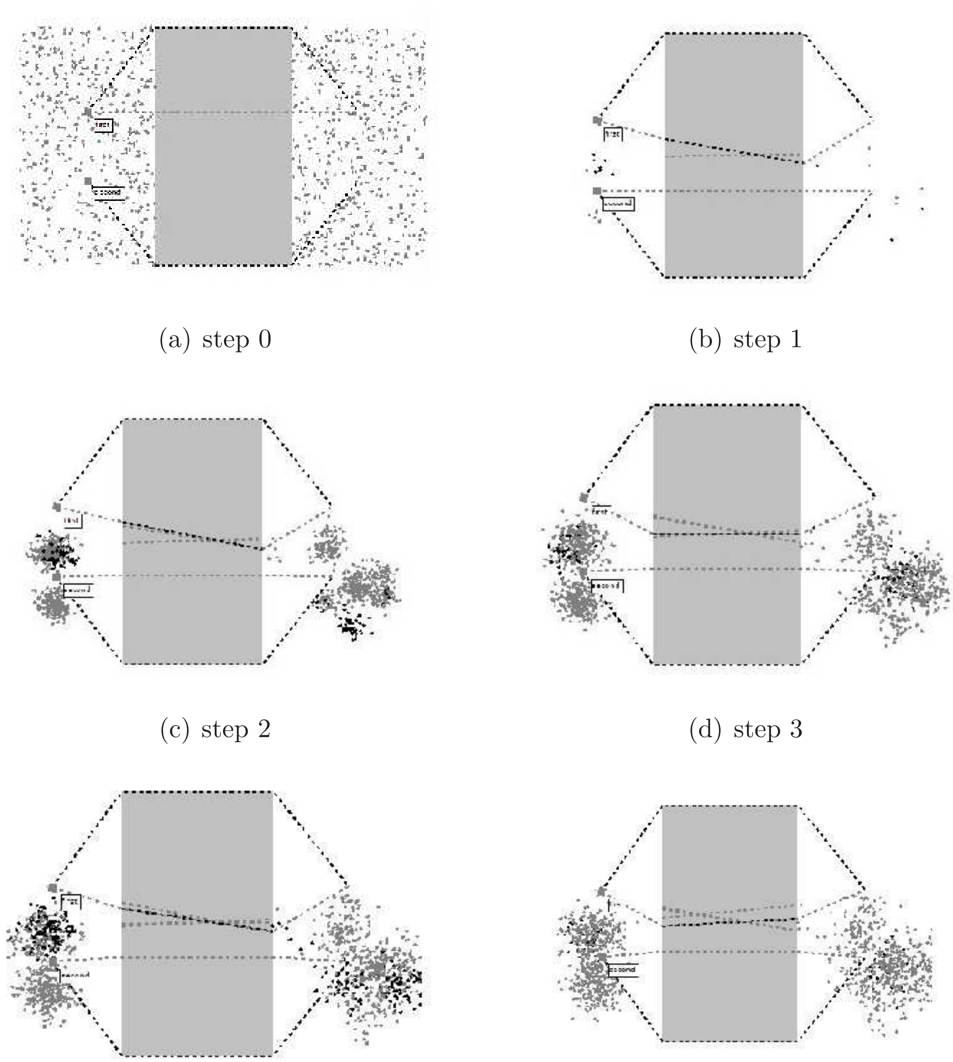
8.5.3 Opponent modeling in the PF agent

We will illustrate the opponent modeling in the PF agent and the overall benefits of the approach by tracing an example where the MCS agent ends in conflict, while a PF agent with the same selfishness and optimism succeeds in negotiating a deal.

Let us consider the negotiation in Figure 8.1-c, where A is an MCS agent with $\lambda = 0.8$ and $\gamma = 1$, which ends in conflict. We repeat the experiment, replacing the MCS agent with a PF agent, with the same λ and γ values.

Figure 8.3 shows the evolution of the particle filter through the first six steps of negotiation. In this figure, a particle is represented by two dots – one on the current location side and one on the destination side. The particles from the cluster chosen for offer formation are shown in black, while the others in gray.

We note that the particles show a relatively large spread which changes from step to step. This is a result of the way in which the offers are formed based on the strongest cluster. If the opponent declines the offer, this represents a strong negative feedback to the selected cluster. This leads to a large variation in the particle cloud, further amplified by the resampling step. Nevertheless, the particle clouds track relatively well the current location



(a) step 0

(b) step 1

(c) step 2

(d) step 3

(e) step 4

(f) step 5

Figure 8.3: Evolution of the opponent model in the PF agent. The black dots and the corresponding dashed line is the cluster selected for offer formation. The gray dots are particles belonging to the other clusters.

and destination of the opponent, which allows the PF agent to choose better offers from the offer pool. In our case, at negotiation round 10, the opponent accepts the PF agents offer, and they move together to their meeting location. Thus, the PF agent, under the same selfishness and optimism parameters, and starting from zero knowledge, could "save" a deal, which was lost for a MCS agent using the same parameters.

8.5.4 Statistical performance comparison

The quality of a specific action strategy / negotiation strategy pair can be measured by the average utility of the deals it can reach over a set of randomly chosen representative scenarios against specific opponents. The statistical averaging is necessary because some strategies might be a better fit for certain scenarios: for instance, fully pessimistic action strategies will yield the best performance in scenarios where no deal is possible.

Figure 8.4 shows the performance of four strategies with various values for optimism and selfishness. The top figure shows the relative utility obtained while the bottom the number of cases where a deal was formed. For all experiments, the opponent uses the MCS strategy with $\lambda = 0.6$ and $\gamma = 1$, The four strategies are MCS, UC and PF to which we add MCSN, a variant of MCS where the action strategy is to not move until a deal is agreed or the negotiation is broken. Thus, the MCSN agent does not perform acting while negotiating, and the optimism parameter has no impact in this case.

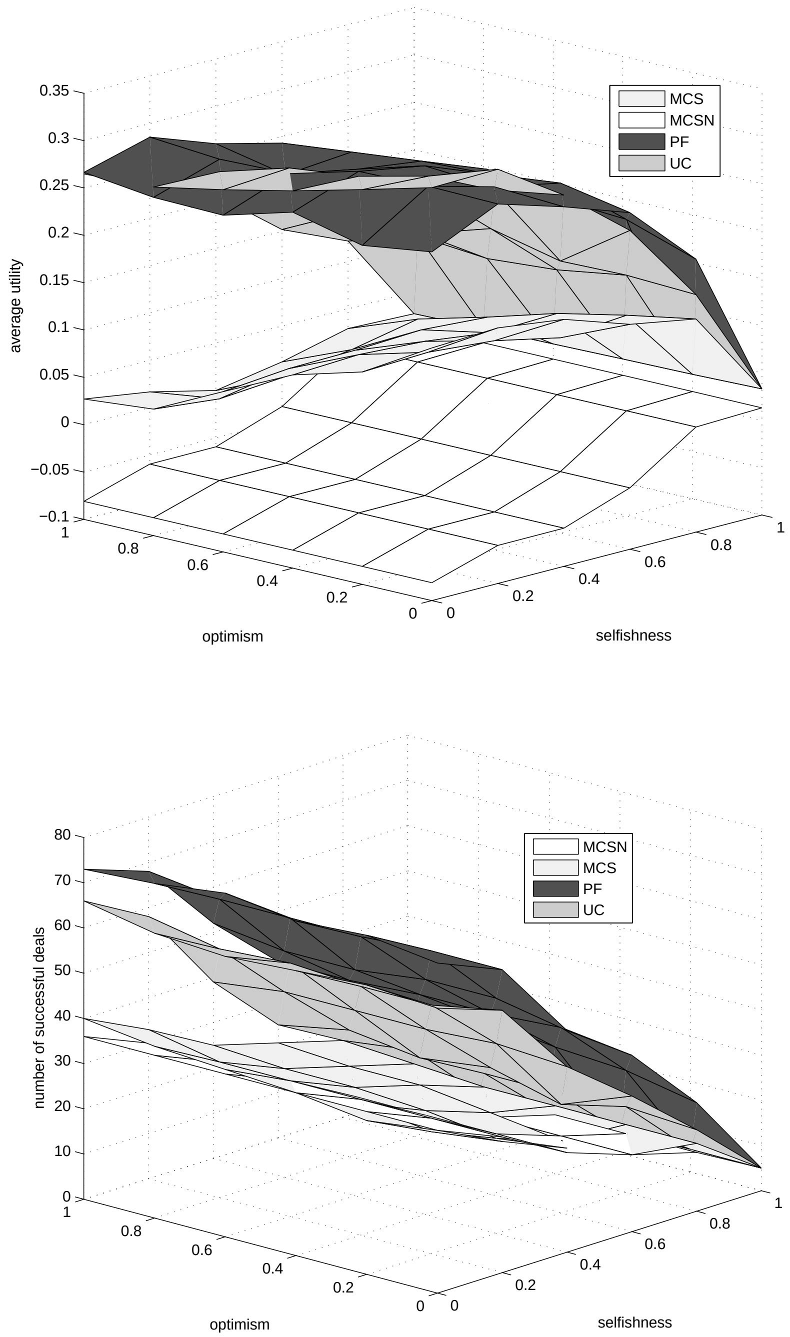


Figure 8.4: Relative performance of various strategies negotiating with (another MCS agent) over 100 scenarios. Top: average utility, bottom: number of successful deals

The first obvious conclusion is that the all the proposed acting while negotiating strategies outperform the MCSN "don't act" strategy. There is also a clear advantage of UC and PF approaches compared to MCS in terms of average utility and number of deals, for every combination of optimism and selfishness. There is a relatively smaller difference between PF and UC. The PF strategy obtains a higher percentage of successful deals and it achieves a higher average utility for the majority of optimism and selfishness values.

Different strategies obtain their maximum utilities at different selfishness and optimism values. For MCS this value is at $\lambda = 0.8$ and $\gamma = 0$. For PF and UC it is around $\lambda = 0.6$ and $\gamma = 0.2$. Note, however, that these values are dependent on the opponent, and further studies with a range of opponents are necessary before definitive conclusions can be drawn.

CHAPTER 9
CONCLUSION

In this dissertation, we considered a class of problems which use negotiation to establish collaboration in space and time. We have shown that the "split the pie" game, frequently used as a canonical problem in studies concerning multi-issue negotiation, can not adequately model this class of problems and we highlighted five characteristics of the class of problems which need to be modelled by a representative canonical problem. We proposed the "Children in the Rectangular Forest" (CRF) model as a canonical problem for the class of applications we are considering, and we have shown that it exhibits the identified characteristics, while keeping the intervening formulas simple and of a low computational complexity.

The CRF problem is the problem of negotiating convoy formation under time constraints. This is a relatively complex multi-issue negotiation: not all the offers are feasible; the utility is a non-linear function of the issues and the offer formation is difficult, as it might require complex path calculations. We introduced a set of metrics which allows us to measure the performance of a negotiation strategy in comparison to its peers, as well as to evaluate, without assuming any particular negotiation strategy whether a negotiation will be easy or difficult. We developed a collaborativeness metric which allows us to put a quantitative value on our intuition of "easy" and "hard" negotiation scenarios. The metric is not dependent on the negotiation strategy, and can be evaluated only by a full knowledge supervisor.

We discussed three negotiation protocols. In the simple exchange of binding offers protocol, the offers are binding for the agents who made them, in the sense that once made by an agent and accepted by the other agent, the offer will be the outcome of the negotiation. The simple protocol can be enhanced by allowing agents to add additional information to aid the negotiation partner in offer formation. We called it exchange of binding offers with mandatory, non-binding evaluations, and we demonstrated that it effectively improves the negotiation performance. In the third protocol, the agents can optionally evaluate the received offers and send back arguments which describe criticisms on the offers which are satisfactory from their own point of view.

We investigated three offer formation strategies. In Monotonic Concession in Space, agents concede towards the opponent's offer in space domain. The Internal Negotiation Deadline sets up a negotiation deadline and lets agents concede evenly towards the opponent's offer in space domain. The Uniform Concession considers the problem in the utility domain, and at each round the agent selects an offer similar to the opponent's offer from a pre-calculated offer pool.

We designed a set of strategies which apply argumentation in the negotiation. The idea is through exchanging arguments, the agents discover their negotiation stage. At some stages, the agent can choose to insist on the previous offer and force the opponent to concede. Finally, by exploring these negotiation protocols, offer formation strategies and argumentation based strategies, we compare the performance of all these strategies. We found the best strategy to be Uniform Concession with Argumentation, because it tends to find the balanced solution from the supervisor point of view.

We described an approach through which agents starting with zero knowledge can estimate the collaborativeness of the scenario using information acquired from the first several negotiation rounds. We showed how this estimation can be used to augment the negotiation strategy with collaborativeness analysis. We demonstrated that the augmented strategies significantly outperform the original strategies for low collaborativeness scenarios and closely match them for high collaborativeness scenarios.

Another approach we investigated was the application of Bayesian learning in the negotiation. The agent can guess the opponent's preference from the sequence of offers it received. We designed three approaches to distinguish probabilities. We showed that using these approaches the agent can gradually identify the opponent and these information can be immediately applied in the next negotiation round which gain some benefits comparing those agents who do not learn.

Another scenario we have investigated involves where agents can act while negotiating. We realized that we need to design the action strategy cooperating with the negotiation strategy. We created a general meta-strategy to control the selfishness and optimism of the agents. We used a particle filter to represent the states and the beliefs about the opponent agent. Through a series of experiments, we demonstrated that the learning agent's beliefs converge to the opponent agent model, and it gains more performances in collaborative scenarios.

The future research can be extended into three directions. First, we can use the beliefs to adapt the agent's selfishness and optimism in acting while negotiating. The weighted particles disclose the information of the scenario and the strategy of the opponent. The agent can apply this to improve the negotiation in next rounds. For example, the agent can increase the selfishness if it believes that the scenario is collaborative. It can be more optimistic if it believes the opponent is "nice". The selfishness and optimism can be updated dynamically as long as the particles are evolved in the agent's mind.

The second direction is the application of Zeuthen's bargaining model in the CRF problem to assist the offer formation strategy, specifically to find out when and how much the agent should concede. The *risk to break the negotiation* in the model can be interpreted as the pragmatic utility minus the baseline utility. For example, the agent who used to be more optimistic is not willing to get the conflict deal as it already moves towards the possible deal. The agent would concede more in the next counter offer.

Finally we can consider a scenario where the learning agent has an opponent who is also learning. In that scenario, the opponent's belief to the agent itself should be added in the particles. It is *the belief about beliefs* or *nested beliefs*. Such assumptions make the spatio-temporal negotiation problems more complex however they will also show a more accurate representation of human behavior in similar situations.

LIST OF REFERENCES

[BKT06] L. Bölöni, M.A. Khan, and D. Turgut. "Agent-based coalition formation in disaster response applications." In *Proceedings of the IEEE Workshop on Distributed Intelligent Systems*, pp. 259–264, June 2006.

[BOR92] K. Binmore, M.J. Osborne, and A. Rubinstein. "Noncooperative models of bargaining." In R. J. Aumann and S. Hart, editors, *Handbook of Game theory with Economic Applications*, volume 1, pp. 179–225. Elsevier, 1992.

[Bra90] S. J. Brams. *Negotiation Games: Applying Game Theory to Bargaining and Arbitration.* New York,NY: Routledge., 1990.

[CM96] David Carmel and Shaul Markovich. "Learning Models of Intelligent Agents." In Howard Shrobe and Ted Senator, editors, *Proceedings of the Thirteenth National Conference on Artificial Intelligence and the Eighth Innovative Applications of Artificial Intelligence Conference, Vol. 2*, pp. 62–67, Menlo Park, California, 1996. AAAI Press.

[CV07] Elisabeth Crawford and Manuela Veloso. "An experts approach to strategy selection in multiagent meeting scheduling." *Autonomous Agents and Multi-Agent Systems*, **15**(1):5–28, 2007.

[DG05] P. Doshi and P.J. Gmytrasiewicz. "Approximating state estimation in multiagent settings using particle filters." In *International Joint Conference on Autonomous Agents and Multi Agent Systems*, 2005.

[Dmy08] Koen Hindriks Dmytro Tykhonov. "Opponent Modelling in Automated Multi-Issue Negotiation Using Bayesian Learning." In *Seventh International Conference on Autonomous Agents and Multiagent Systems (AAMAS 08)*, pp. 331–338, 2008.

[End06] Ulle Endriss. "Monotonic Concession Protocols for Multilateral Negotiation." In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, 2006.

[FP08] S.G. Ficici and A Pfeffer. "Simultaneously Modeling Humans' Preferences and their Beliefs about Others' Preferences." In *Seventh International Conference on Autonomous Agents and Multiagent Systems (AAMAS 08)*, pp. 323–330, 2008.

[FWJ02] S.S. Fatima, M. Wooldridge, and N.R. Jennings. “Multi-issue negotiation under time constraints.” In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1*, pp. 143–150. ACM New York, NY, USA, 2002.

[FWJ06] S. Shaheen Fatima, Michael Wooldridge, and Nicholas R. Jennings. “Multi-Issue Negotiation with Deadlines.” *J. Artif. Intell. Res. (JAIR)*, **27**:381–417, 2006.

[GA04] A. Garland and R. Alterman. “Autonomous agents that learn to better coordinate.” *Autonomous Agents and Multi-Agent Systems*, **8**(3):267–301, 2004.

[GD97] M. Golfarelli and S. Rizzi D. Maio. “Multi-Agent Path Planning Based on Task-Swap Negotiation.” In *Proceedings 16th UK Planning and Scheduling SIG Workshop*, pp. 69–82, 1997.

[GD05] P.J. Gmytrasiewicz and P. Doshi. “A framework for sequential planning in multi-agent settings.” *Journal of Artificial Intelligence Research*, **24**(1):49–79, 2005.

[GR00] M. Golfarelli and S. Rizzi. “Spatio-temporal clustering of tasks for swap-based negotiation protocols in multi-agent systems.” In *Proceedings 6th International Conference on Intelligent Autonomous Systems*, pp. 172–179, 2000.

[HPC93] Steve Hanks, Martha E. Pollack, and Paul R. Cohen. “Benchmarks, Test Beds, Controlled Experimentation, and the Design of Agent Architectures.” *AI Magazine*, **14**(4):17–42, 1993.

[IKH08] T. Ito, M. Klein, and H. Hattori. “A multi-issue negotiation protocol among agents with nonlinear utility functions.” *Multiagent and Grid Systems*, **4**(1):67–83, 2008.

[JFL01] N. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, C. Sierra, and M. Wooldridge. “Automated negotiation: prospects, methods and challenges.” *International Journal of Group Decision and Negotiation*, **10**(2):199–215, 2001.

[KB06] M. A. Khan and L. Bölöni. “Negotiation-based coalitions in the physical world.” In P. Stone and G. Weiss, editors, *Fifth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-2006)*, pp. 411–413, May 2006.

[KLM96] L.P. Kaelbling, M.L. Littman, and A.W. Moore. “Reinforcement learning: A survey.” *Arxiv preprint cs/9605103*, 1996.

[KM06] A. Kakas and P. Moraitis. “Adaptive agent negotiation via argumentation.” In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, p. 391. ACM, 2006.

[LB07] Y. Luo and L. Bölöni. "Children in the forest: towards a canonical problem of spatio-temporal collaboration." In *The Sixth Intl. Joint Conf. on Autonomous Agents and Multi-Agent Systems (AAMAS 07)*, pp. 986–993, 2007.

[LB08a] Y. Luo and L. Bölöni. "Collaborative and competitive scenarios in spatio-temporal negotiation with agents of bounded rationality." In *Proceedings of the 1st International Workshop on Agent-based Complex Automated Negotiations , in conjunction with the The Seventh Intl. Joint Conf. on Autonomous Agents and Multi-Agent Systems (AAMAS 08)*, pp. 40–47, 2008.

[LB08b] Y. Luo and L. Bölöni. "Learning models of the negotiation partner in spatio-temporal collaboration." In *The 4th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom-2008)*, November 2008.

[LB09a] Y. Luo and L. Bölöni. "Acting while negotiating in the convoy formation problem." In *The Second International Workshop on. Agent-based Complex Automated Negotiations (ACAN 2009)*, May 2009.

[LB09b] Y. Luo and L. Bölöni. "Analyzing and exploiting the competitiveness of scenarios for negotiating convoy formation under time constraints." *to be published in Multiagent and Grid Systems - an International Journal*, 2009. Special Issue of Advances in Agent-mediated Automated Negotiations.

[LC04] Jian Li and Yuan-Da Cao. "Bayesian Learning in Bilateral Multi-Issue Negotiation and Its Application in MAS-Based Electronic Commerce." *iat*, **00**:437–440, 2004.

[OR94] M. J. Osborne and A. Rubinstein. *A Course in Game Theory.* The MIT Press, 1994.

[PL05] L. Panait and S. Luke. "Cooperative multi-agent learning: The state of the art." *Autonomous Agents and Multi-Agent Systems*, **11**(3):387–434, 2005.

[RRJ03] Iyad Rahwan, Sarvapali D. Ramchurn, Nicholas R. Jennings, Peter McBurney, Simon Parsons, and Liz Sonenberg. "Argumentation-Based Negotiation.", 2003.

[Rub82] A. Rubinstein. "Perfect equilibrium in a bargaining model." *Econometrica*, **50**(1):97–109, 1982.

[SS06] Sabyasachi Saha and Sandip Sen. "Negotiating efficient outcomes over multiple issues." In *5th International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2006)*, pp. 423–425, 2006.

[SV99] Tuomas Sandholm and Nir Vulkan. "Bargaining with Deadlines." In *AAAI/IAAI*, pp. 44–51, 1999.

[Tan93] M. Tan. “Multi-agent reinforcement learning: Independent vs. cooperative agents.” In *Proceedings of the tenth international conference on machine learning*, volume 337, 1993.

[WJ94] Michael J. Wooldridge and Nicholas R. Jennings. “Agent Theories, Architectures, and Languages: A Survey.” pp. 1–39, 1994.

[You75] O. R. Young. *Bargaining : Formal Theories of Negotiation.* Urbana, IL: University of Illinois Press, 1975.

[ZS98] Dajun Zeng and Katia Sycara. “Bayesian learning in negotiation.” *Int. J. Hum.-Comput. Stud.*, **48**(1):125–141, 1998.