Enacting Protocols by Commitment Concession*

Pınar Yolum
Department of Computer Engineering
Boğaziçi University
Bebek, Istanbul, TR-34342, Turkey
pinar.yolum@boun.edu.tr

Munindar P. Singh
Department of Computer Science
North Carolina State University
Raleigh, NC 27695-7535, USA
singh@ncu.edu

ABSTRACT

Commitment protocols formalize interactions among autonomous, heterogeneous agents, leaving the agents' local policies unspecified. This paper studies the problem of agents enacting commitment protocols, which inherently requires that their policies cohere with the given protocols. Specifically, in many important settings, if agents incautiously create and discharge commitments, they can expose themselves to certain risk; conversely, if the agents are (excessively) cautious, a protocol enactment may deadlock. This paper adopts the well-known idea of monotonic concession, but specializes and enhances it with the particular features of commitments. Specifically, this paper formulates inference rules for commitment concession that respect the nature of commitments. Next, it shows how commitments can be systematically revised as the agents incrementally engage each other in enacting their protocol. This paper demonstrates how such rules can be applied in practice, and identifies conditions under which progress and termination of protocol enactment can be guaranteed.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent Systems

General Terms

Theory; algorithms

Keywords

Communication: protocols; argumentation

1. INTRODUCTION

Protocols regulate interactions between agents. Commitments are elements of contractual relationships among autonomous parties [9]. In this manner they provide a basis for modeling the "content" of the interactions among agents without regard to the lower level operational details of how messages are exchanged. As a result, commitments support improved autonomy and heterogeneity for the agents constituting a multiagent system.

Several researchers recognize that traditional formalisms used in modeling network protocols, such as finite state machines (FSMs) and Petri Nets, specify protocols merely in terms of legal sequences or concurrent combinations of actions without regard to the meanings of those actions. When directly applied to multiagent settings, the above approaches lead to protocols that are over-constrained [11, 12, 7]. Consequently, commitment protocols have become quite prevalent in recent years.

Like all interaction protocols, commitment protocols specify how their participants may interact. They leave the important aspect of strategy or *local policy* to the agents and, presumably, to the principals of the agents. In other words, a commitment protocol specifies what messages are allowed and the consequences of the messages on the evolving state of a protocol being enacted. But a protocol does not specify what messages to send.

Since commitments (along with domain propositions) offer a declarative semantics characterizing the evolving state of the enactment of a protocol, they provide a basis for the participating agents to reason about their actions. Moreover, the declarative semantics provides a principled basis for agents to interleave contextually relevant, but previously unspecified, actions into the enactment of a protocol.

Previous work has studied scenarios involving commitment protocols to demonstrate their flexibility. However, some important questions underlying the successful *enactment* of protocols are not adequately addressed by previous work. As our running example, we use the well-known purchase protocol. Using this example helps us explain our approach in comparison to previous work without having to introduce irrelevant features.

EXAMPLE 1. *A customer and a merchant participate in a purchase transaction where they exchange a goods for money. For this transaction to succeed, the customer has to pay and the seller has to deliver the goods. This business transaction can be modeled as a commitment protocol. If the customer sends a payment before receiving the goods from the seller, then he will be at risk until the seller delivers the goods. Similarly, if the seller sends the goods before receiving the payment, then he is at risk until the customer sends the payment.* ∎

In other words, progress in enacting a protocol thus depends upon some agent "venturing out" and doing its part before other agents do their parts. Commitment protocols often involve condi-

*The first author is partially supported by Boğaziçi University Research Fund under grant BAP06A103 and The Scientific and Technological Research Council of Turkey by a CAREER Award under grant 105E073. The second author's research was partially supported by the US National Science Foundation under grant DST-0139037. We thank the reviewers for helpful comments.

978-81-904262-7-5 (RPS) ©2007 IFAAMAS

tional commitments (formalized below). Often, especially in two-party settings, such commitments refer to each others' conditions in a complementary manner. The following is a telling example:

EXAMPLE 2. *The merchant commits to sending the goods only if the customer pays the money. The customer commits to making the payment if the merchant sends the goods. In a naïve operationalization, the merchant waits for the customer to pay, while the customer waits for the merchant to deliver the goods. Thus deadlock ensues.* ▌

Agents enact a protocol by making commitments to each other and manipulating these commitments as they see fit. However, in settings where agents lack trust in each other, making strong commitments up front is not desirable since other agents may not make corresponding commitments or may not fulfill their commitments.

To resolve the problem depicted in Example 1, we can possibly employ an approach based on the *monotonic concession protocol*, a well-known approach to negotiation (the term "protocol" above is used somewhat differently from the usage of this paper) [3, 8]. Concession is based on the idea of the agents involved incrementally conceding to one another by proposing deals that are presumably more desirable to the others involved. A negotiation succeeds if all parties agree on a deal. In the present setting, concession based on commitments is considered. Commitment concession is seamlessly incorporated within a commitment protocol enactment.

Although traditionally concession is understood generically, in the present setting, we can take advantage of the structure of commitments to systematically specify the kinds of concessions that might be appropriate. Further we can express possible concessions qualitatively as inference rules.

The main idea of commitment concession is that the agents incrementally commit themselves at each round, increasing the "risk" that they take as the other parties also increase their risks. In other words, participants make weak commitments in the beginning and revise or discharge these commitments as other agents revise or discharge their commitments. Hence, agents commit themselves via a series of concessions. Before making a move, each participant needs to calculate the consequence of its move in terms of the move's risks and benefits.

Each agent can keep track of the risks and benefits of its actions. As a result, an agent can decide if other participants in the protocol are being cooperative, and making suitable commitments. If so, the agent can revise its commitment and make a stronger commitment. Otherwise, the agent can stop enacting the protocol with as little loss as possible.

The contributions of this paper are as follows. This paper

- Tackles the problem of enacting commitment protocols in a way that respects the policies of the participating agents.
- Develops a qualitative notion of risk and benefit based on commitments.
- Shows how a range of concessions can be constructed based on commitments, and expressed as inference rules.
- Establishes technical properties of the above sets of inference rules.

The rest of this paper is organized as follows: Section 2 provides background on commitments, protocols, and concession. Section 3 explains how agents can calculate the risks and benefits of an action to decide if they should revise their commitment. Section 4 defines commitment concession rules formally, gives examples of concessions, and develops their properties. Section 5 discusses our work with reference to the literature.

2. TECHNICAL BACKGROUND

Commitments are directed from one agent to another. In essence, the debtor of a commitment is obliged to bring about a specified condition or proposition. Commitments result from communicative actions. That is, agents create commitments and manipulate them through the protocol they follow. We represent commitments themselves as propositions. Previous research has represented commitments and operations on them in notations such as the event calculus [12]. The notation is not central to the claims of this paper, but can be assumed to be the event calculus for concreteness.

DEFINITION 1. $\mathsf{C}(x, y, p)$ *is a base-level commitment of* debtor x *to* creditor y *to bring about proposition* p *[9].*

When a commitment of this form is created, intuitively, the effect is that the debtor becomes responsible to the creditor for satisfying p. The commitment would be discharged if and only if p holds sometime in the future. Note that the proposition p does not involve other commitments.

DEFINITION 2. $\mathsf{CC}(x, y, p, q)$ *is a* conditional commitment*: if proposition* p *is satisfied,* x *will become committed to bringing about proposition* q*. Both* p *and* q *can refer to other commitments. Here* p *is considered the* condition *of this commitment.*

Conditional commitments are useful when a party wants to commit only if a certain condition holds. Such a condition would typically be brought about directly or indirectly by the other party. Further, such a condition can involve the discharge of another commitment. Below, *goods* means that goods are delivered and *pay* means that payment is made.

EXAMPLE 3. *A base-level commitment C(merchant, customer, goods) means that the merchant commits to the customer to have the goods delivered. A conditional commitment, CC(merchant, customer, pay, goods) specifies that the merchant will commit to sending the goods if the customer pays.* ▌

Singh introduced six operations to create and manipulate commitments [9]: of these, for simplicity, this paper considers only the following. Below, x and y are agents and c is a commitment.

- *Create(x, c)* establishes the commitment c. This can only be performed by the debtor of the commitment x.
- *Discharge(x, c)* resolves the commitment c. This can only be performed by the debtor of the commitment to mean that the commitment has successfully been carried out. Discharging a commitment terminates that commitment.

A base-level commitment $\mathsf{C}(x, y, p)$ is discharged when its proposition p is achieved. The base-level commitment is then no longer in force. When the condition of a conditional commitment holds, a commitment for its proposition is created and the conditional commitment itself is no longer in force. If the proposition of a conditional commitment occurs, then the conditional commitment is considered immediately discharged. Thus p and $\mathsf{C}(x, y, p)$ cannot coexist. Likewise, q and $\mathsf{CC}(x, y, p, q)$ cannot coexist. Further, p and $\mathsf{CC}(x, y, p, q)$ cannot coexist, but the occurrence of p causes the creation of $\mathsf{C}(x, y, q)$.

There are thus two main ways in which a commitment $\mathsf{C}(x, y, q)$ may be created. Either agent x creates $\mathsf{C}(x, y, q)$ directly or a conditional commitment of the form $\mathsf{C}(x, y, p, q)$ already exists and somehow p is brought about. We call this last operation *detach*. Typically, but not necessarily, agent y would bring about p.

2.1 Specifying Commitment Protocols

A *protocol specification* is based on a set of roles and a vocabulary of messages understood as communicative actions performed by (agents playing) the various roles. Importantly, for reasons well-known in the literature [11], the messages can be given a meaning in terms of the conditions they bring about and the operations on commitments to which they correspond.

Given a set of atomic propositions $\mathcal{P}$ and a set of roles $\mathcal{R}$, we can generate propositions as follows: (1) the set of Boolean combinations of members of $\mathcal{P}$; (2) the set of base-level commitments $\mathcal{B}$ from (1); and the set of conditional commitments $\mathcal{C}$ from $\mathcal{B}$ and $\mathcal{P}$. The commitments are simply expressions involving special operators C and CC, which take two distinct roles from $\mathcal{R}$ and, respectively, one proposition or two propositions. These form the logical expressions that we deal with.

A commitment protocol P is defined over the above by specifying a set of messages and how they affect the state of the protocol (expressed in terms of the above logical expressions). The semantics of the messages can be specified in a suitable formalism, such as the event calculus [11, 12]. The choice of formalism is not critical for the present paper. The various action formalisms yield transition systems whose states are the states of the world and whose transitions are the messages being exchanged. The initial and final states, or transitions are not explicitly specified: these are inferred from the protocol specification to produce the transition system. States in which there are pending commitments are not allowed to be final, since the existence of an undischarged commitment indicates some kind of an exceptional state.

2.2 Monotonic Concession

Monotonic concession enables negotiating parties to reach agreement by iteratively revising their offers [8]. Initially, two agents may state offers that are not acceptable to the other party. At each round, both agents incrementally make a move toward what can be acceptable to the other party. There are two obvious outcomes of monotonic concession. One possibility is that both agents arrive at a mutually acceptable deal and thus reach an agreement. The second possibility is that for both of the agents making another compromise from the current offer is unacceptable: hence a conflict occurs and agreement is not reached.

Intuitively, such incremental concession is particularly useful for minimizing risks in interactions. At each round of concessions, both agents increase the risk they are taking for the sake of reaching an agreement. But because each agent's increase in risk is countered by a potential increase in risk taken by other agents, the net risk thus would not increase abruptly.

3. RATIONAL BASIS FOR ENACTMENT

The basic intuition underlying the operationalization of commitment protocols is that (1) the current state of the protocol (termed the *protocol state*) is maintained during enactment; (2) at that state, the relevant propositions (including domain propositions and commitments) are true or false; (3) based on the truth and falsity of those propositions, the preconditions of some actions can be satisfied; (4) any of the actions whose preconditions are satisfied can be performed by the appropriate role; (5) this action causes the current state to evolve; (6) the protocol terminates when a final state is reached. For simplicity, we assume that only one action can happen at one time point. Hence, we are not concerned with concurrent actions. Also, like the works cited above, we finesse the challenges of message delay. Additional subtleties of enactment are the theme of this paper and are discussed below.

3.1 States of Protocol Enactment

A state is modeled as a set of propositions that hold in it. Each agent values a state based on its propositions. As explained in Section 2.1, three kinds of propositions are relevant here: atomic propositions, and base-level and conditional commitments.

DEFINITION 3. *Given a commitment protocol T, the* space of states *for T, $\mathcal{S}$, is given by maximal consistent sets of propositions in $\mathcal{P} \cup \mathcal{B} \cup \mathcal{C}$.*

Based on the above, a state can be represented exactly by the subset of $\mathcal{P} \cup \mathcal{B} \cup \mathcal{C}$ whose members are true in that state. The propositions not in the set are interpreted as being false.

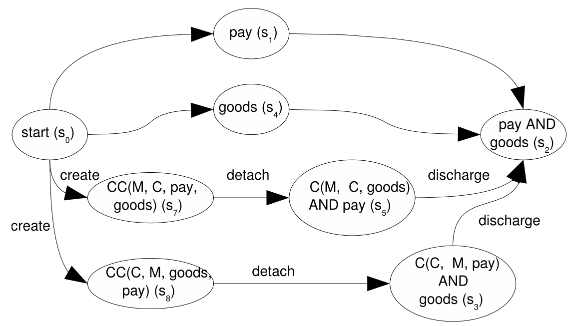

Figure 1: Some example enactments of a commitment protocol for purchasing

Let's now consider how the purchase protocol might be enacted. Figure 1 depicts some of its possible enactments. The vertices represent protocol states. Each state is identified by the propositions and commitments that hold in it: that is, two states must have some distinctions. Each edge label denotes the commitment operation for the corresponding transition.

Base-level commitments rarely exist in isolation. They are usually part of a series of protocol steps that are carried out to realize a goal state. One way to think about them is based on turn-taking. When each agent gets a turn, it can create a commitment and (subsequently) discharge it.

EXAMPLE 4. *Following Example 3, the merchant wants to sell the goods in exchange for payment. With turn-taking, the merchant and customer can perform the following operations (in order):*

1. create(merchant, C(merchant, customer, goods))
2. discharge(merchant, C(merchant, customer, goods))
3. create(customer, C(customer, merchant, pay))
4. discharge(customer, C(customer, merchant, pay)) ▮

3.2 Understanding Benefits and Risks

A rational agent would consider the *benefit* and the *risk* of any commitment it creates. The benefit of a commitment is what the agent will gain by creating the commitment. For example, the agent may cause another party to become committed to it (because of an existing conditional commitment that it exercises). The risk of the commitment is what the agent will have to do to discharge the commitment, which it can potentially lose if the other agents do not deliver as expected.

Table 1: Risks and benefits to customer (C) of C's commitments

Commitment made	Risk	Benefit
CC(C, M, goods, pay)	C(C, M, pay)	goods
CC(C, M, C(M, C, goods), pay)	C(C, M, pay)	C(M, C, goods)
C(C, M, pay)	pay	None

Risks and benefits are summed over all the commitments. For example, if $\mathsf{C}(x, z, p, q)$ and $\mathsf{C}(y, z, p, r)$ hold, then the benefit to z of p is the sum of the benefits of q and r.

Each participant in the protocol assesses the risks and benefits of each commitment from its point of view. When calculating the risks and benefits, it is important to decide how long into the future the agent is willing to look. For example, the creation of a commitment may not have an immediate benefit, but if it will force the other agent to make a commitment in the near future, it may still be considered beneficial. For now, we assume that the agent looks only one state ahead in the future. Table 1 lists the risks and benefits of some commitments from the customer's point of view.

The risk and benefit of a commitment are duals of each other. Each agent is interested in creating a commitment where its risk is minimized and its benefit is maximized. Minimizing the risk means that the debtor's share will be little, whereas maximizing the benefit means that the gain from the commitment will be high.

EXAMPLE 5. *Consider the examples of Table 1.*

- *For CC(C, M, goods, pay), the customer's benefit is greater than its risk, since when the merchant accepts this commitment it will deliver the goods whereas the customer will only be committed to paying after it receives the goods. Hence, if the merchant never delivers the goods, then the customer has no obligation to pay.*

- *For CC(C, M, C(M, C, goods), pay), if the merchant accepts the conditional commitment, both parties become committed to each other: the merchant to delivering the goods and the customer to paying. However, the customer has committed to paying even if the merchant does not deliver the goods. Compared to the first case, the benefits of this commitment are worse than those of CC(C, M, goods, pay).*

- *For C(C, M, pay), the commitment has no benefit for the customer but a risk of paying. Suppose the customer makes such a commitment hoping that the merchant will also create a commitment for delivering goods. Then if the merchants does not create its commitment, the customer will be at a disadvantage.*

Overall, when these three commitments are compared, the customer is best off making the first commitment, since its benefits clearly outweigh its risks. ∎

In a two-party protocol, the risks and benefits of the participants are correlated with each other. For example, to consider the merchant's point of view on risks and benefits, it is enough to flip the second and third columns in Table 1. Thus, the second column shows the merchant's benefit and the third column the merchant's risk.

3.3 Costs and Valuations

Each agent has a valuation for each state based on the propositions that hold therein. We distinguish between task and goal propositions. An agent's valuation of its tasks is negative; its valuation of its goals is positive. The valuations of an agent do not change during the enactment of the protocol. Formally, valuations are defined for states (sets of propositions), but are applied to individual propositions for simplicity. We write $v_x(\cdot)$ to express agent x's valuation.

For informal motivation, it helps to consider the *credits* and *debts* of an agent based on the commitments that hold. These correspond to positive and negative valuations, respectively. If $v_x(p) > 0$ and $\mathsf{C}(y, x, p)$, then this is a credit for x. If $v_x(p) < 0$, then $\mathsf{C}(y, x, p)$ would be a threat from y. Following Castelfranchi's dictum that commitments be desirable for their creditors [2], we remove such commitments. If $v_x(p) < 0$ and $\mathsf{C}(x, y, p)$, then this is a debt of x. If $v_x(p) > 0$, then $\mathsf{C}(x, y, p)$ would be a self-serving commitment. We do not have to eliminate such commitments, because they would occur in win-win situations such as teamwork. Alice commits to Bob to bringing about the condition where their team wins, but Alice happens to benefit from having the team win.

The valuations given by each agent to an atomic proposition are arbitrary. This reflects the important intuition behind independent private valuations of economic theory. In general, agents have different valuations of the same state: this is what makes trade among agents individually rational for each of them. For example, *customer* may value *pay* (paying \$10) at $-\$10$ and (receiving the) *goods* at $+\$11$, whereas *merchant* may value *pay* (getting paid \$10) \$10 at $+\$10$ and (sending the) *goods* at $-\$9$. Thus both would see a gain of \$1 from carrying out the trade.

The valuation given to a base-level commitment has a magnitude no larger than the magnitude of the valuation of its condition. For a reputable agent x and a condition p with a positive valuation by agent y, a base-level commitment $\mathsf{C}(x, y, p)$ is given a nonzero valuation by y. This captures that x is trusted by y.

DEFINITION 4. *A valuation function* $v_x : \mathcal{S} \mapsto (-\infty, \infty)$ *is* coherent *if and only if the following conditions hold:*

Null. *Valuation of an empty set is zero, because there is nothing to deliver or expect:* $v_x(\{\ \}) = 0$

Separability. *Valuation of a union of two sets is the sum of their valuations:* $v_x(S_1 \cup S_2) = v_x(S_1) + v_x(S_2)$

As creditor. *Commitment for goal is worth less than the deed:* $v_x(p) > 0$ *implies* $0 \leq v_x(\mathsf{C}(y, x, p)) \leq v_x(p)$

As debtor. *Commitment for task is worth more than the deed:* $v_x(p) < 0$ *implies* $0 \geq v_x(\mathsf{C}(x, y, p)) \geq v_x(p)$

As conditional creditor. *Value to creditor of conditional commitment:* $v_x(\mathsf{C}(y, x, p)) \geq v_x(\mathsf{CC}(y, x, q, p))$ $\geq v_x(q) + v_x(\mathsf{C}(y, x, p))$

As conditional debtor. *Value to debtor of conditional commitment:* $v_x(\mathsf{C}(x, y, q)) \leq v_x(\mathsf{CC}(x, y, p, q))$ $\leq v_x(p) + v_x(\mathsf{C}(x, y, q))$

If x is perfectly honest and competent, $v_x(p) = v_x(\mathsf{C}(x, y, p))$, for tasks it has to perform. If $v_x(p) = v_x(\mathsf{C}(x, y, p))$, then the two states are indistinguishable for x in terms of valuations. In most practical settings, the inequality is strict: $v_x(p) > v_x(\mathsf{C}(x, y, p))$.

For an agent (x), being the creditor of a conditional commitment is less preferable than being the creditor of a base-level commitment, since for the latter it need not fulfill any tasks. Hence, $v_x(\mathsf{C}(y, x, p)) \geq v_x(\mathsf{CC}(y, x, q, p))$. But, being a creditor of a conditional commitment is better than having already achieved the condition of the conditional commitment (and thus becoming the creditor of a base-level commitment), since if the base-level commitment is canceled, there is nothing the creditor can do to negotiate. Hence, $v_x(\mathsf{CC}(y,x,q,p)) \geq v_x(q) + v_x(\mathsf{C}(y,x,p))$. The reverse argument holds for the conditional debtor.

As described above, in general, agents have different valuations for the same state. Considering the purchasing protocol, $v_C(pay)+v_C(goods) > 0$ as well as $v_M(pay)+v_M(goods) > 0$. Thus, both parties have incentives for carrying out the protocol and ending at the final state. If the valuation of the final state were not larger that of the initial state, it would not be rational for them to carry out the protocol. Thus, the final state must have strictly higher valuation for each agent than the initial state: $v_x(\mathit{final}) > v_x(\mathit{initial})$.

DEFINITION 5. Social welfare *of a state, $w(s)$, is the sum of the valuations of the state by all participating agents.*

The social welfare of a state measures the benefits of a state when all agents are considered. An agent will not know the social welfare of a state since it does not have access to the valuations of other agents. However, as a whole, one would expect the protocol to help agents come closer to their goals, thus increase the social welfare of the agents over time.

When an agent begins to enact a protocol, it decides on a goal state. The participants of a protocol may have different or even conflicting goal states. Informally, a goal of an agent captures the states of the world that the agent is predisposed to achieve. Formally, a goal is represented by a modal expression similar to commitments.

DEFINITION 6. *Let $p \in \mathcal{S}$. $\mathsf{G}(x,p)$ means that agent x has proposition p as a goal.*

Each agent has fixed goals during the enactment of the protocol and plans its actions to achieve these goals. However, in general, an individual agent cannot form complete plans since other participants' actions influence the outcomes, including whether its goals are achieved. This necessitates a method for agents to reach agreement about how they will carry out their interactions.

4. COMMITMENT CONCESSION

Now we formulate the rules by which agents can concede based on their commitments to each other, and study the properties of these rules. Note that in a traditional concession protocol, agents act at the same time but see their actions only after they have all made their moves. However, here we assume that agents take turns.

4.1 Rules for Concession

When an agent decides to revise its commitments, it can use the following rules to revise its commitments. Inference rules relate antecedents to consequents. The inference rules below are to be interpreted as decision-making rules. If the agent, in fact, takes a decision according to one of these rules, the state of enactment of the protocol would have evolved to another state. In other words, the rules below reflect the actions of exactly one agent during a round of concessions. Concession moves, including those described by these rules are merely allowed: they are not required. Some rules are clearly stronger than others. However, they are included because agents need not respect all or any of the rules.

$$\frac{\mathsf{G}(x,p)}{\mathsf{CC}(x,y,p,q)} \qquad \text{(create-CC)}$$

When starting a negotiation, x should not right away make a base-level commitment, but instead ask for some benefits from y. Intuitively, this leads x to make a conditional commitment towards y in which y brings about a goal proposition for x in return for becoming the creditor of a commitment.

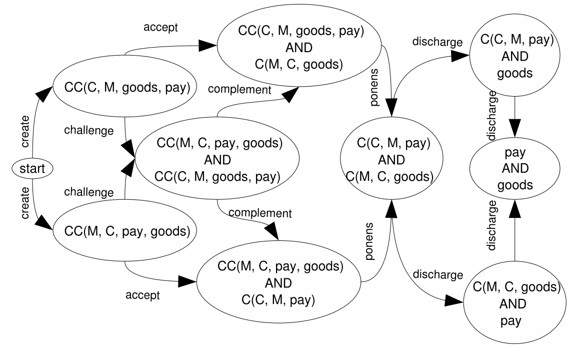

Figure 2: Protocol enactment with commitment concession

$$\frac{\mathsf{C}(x,y,q) \qquad \mathsf{C}(y,x,p) \qquad \mathsf{G}(x,p)}{q} \qquad \text{(discharge-C)}$$

x should consider discharging its commitment only after it guarantees a certain benefit from y, so that it is not left at a disadvantage. In other words, x should only discharge its commitment, after y commits to fulfilling a goal of x.

$$\frac{\mathsf{CC}(y,x,q,p) \qquad \mathsf{G}(x,p)}{\mathsf{C}(x,y,q)} \qquad \text{(accept)}$$

When the conditional commitment $\mathsf{CC}(y,x,q,p)$ exists and when x considers p to be a goal state, it is better for x to commit to $\mathsf{C}(x,y,q)$ than to bring about q immediately. This is because creating $\mathsf{C}(x,y,q)$ has smaller risks for x than bringing about q. In our running example, the customer may say that it is willing to pay if the merchant delivers the goods. Instead of delivering the goods right away, the merchant can *commit to* delivering the goods, thereby making a small step toward its goal state, and simultaneously helping the customer make a small step toward the customer's goal state.

Exercising accept makes sense only when x trusts y to cooperate, since after the execution of this rule, for x the new protocol state has higher risk than benefit. If x lacks sufficient trust in y, it can instead challenge y to make a move.

$$\frac{\mathsf{CC}(y,x,q,p) \qquad \mathsf{G}(x,p)}{\mathsf{CC}(x,y,p,q)} \qquad \text{(challenge)}$$

Following the running example, the customer first states that it is willing to pay if the merchant delivers the goods. If the merchant does not trust the customer to pay after it delivers the goods, it may ask the customer to pay first and *commit to* delivering the goods afterward.

$$\frac{\mathsf{CC}(x,y,p,q) \qquad \mathsf{CC}(y,x,q,p)}{\mathsf{C}(x,y,q) \qquad \neg\mathsf{CC}(x,y,p,q)} \qquad \text{(complement)}$$

Before this rule is executed, two agents have complementary conditional commitments toward each other. In terms of our running example, the customer is willing to pay if the merchant delivers and the merchant is willing to deliver if the customer pays. From both agents' viewpoints, their distance to the goal state is the same. However, if neither one takes a risk, then the protocol cannot progress. Hence, this rule states that if two agents have the same distance to the goal state, then one of the agents will revise its commitment to decrease the distance.

$$\frac{\mathsf{C}(x,y,q) \qquad \mathsf{CC}(y,x,q,p)}{\mathsf{C}(y,x,p) \qquad \neg\mathsf{CC}(y,x,q,p)} \qquad \text{(ponens)}$$

Similarly, when the conditional commitment $\mathsf{CC}(y,x,q,p)$ and the base-level commitment $\mathsf{C}(x,y,q)$ hold, it is better for y to revise its initial commitment and create $\mathsf{C}(y,x,p)$. Following our

Table 2: Valuation changes for x induced by various concession rules when applied by x

Rule	v_x before	v_x after	v_x after $-$ v_x before
create-CC	0	$v_x(\mathsf{CC}(x,y,p,q))$	$v_x(\mathsf{CC}(x,y,p,q)) \leq 0$
challenge	$v_x(\mathsf{CC}(y,x,q,p))$	$v_x(\mathsf{CC}(x,y,p,q)) + v_x(\mathsf{CC}(y,x,q,p))$	$v_x(\mathsf{CC}(x,y,p,q)) \leq 0$
complement	$v_x(\mathsf{CC}(x,y,p,q)) + v_x(\mathsf{CC}(y,x,q,p))$	$v_x(\mathsf{C}(x,y,q)) + v_x(\mathsf{CC}(y,x,q,p))$	$v_x(\mathsf{C}(x,y,q))$ $-v_x(\mathsf{CC}(x,y,p,q)) \leq 0$
accept	$v_x(\mathsf{CC}(y,x,q,p))$	$v_x(\mathsf{CC}(y,x,q,p)) + v_x(\mathsf{C}(x,y,q))$	$v_x(\mathsf{C}(x,y,q)) \leq 0$
ponens	$v_x(\mathsf{C}(y,x,p)) + v_x(\mathsf{CC}(x,y,p,q))$	$v_x(\mathsf{C}(y,x,p)) + v_x(\mathsf{C}(x,y,q))$	$v_x(\mathsf{C}(x,y,q))$ $-v_x(\mathsf{CC}(x,y,p,q)) \leq 0$
discharge-C	$v_x(\mathsf{C}(y,x,p)) + v_x(\mathsf{C}(x,y,q))$	$v_x(\mathsf{C}(y,x,p)) + v_x(q)$	$v_x(q) - v_x(\mathsf{C}(x,y,q)) \leq 0$

Table 3: Valuation changes for y induced by various concession rules when applied by x

Rule	v_y before	v_y after	v_y after $-$ v_y before
create-CC	0	$v_y(\mathsf{CC}(x,y,p,q))$	$v_y(\mathsf{CC}(x,y,p,q)) \geq 0$
challenge	$v_y(\mathsf{CC}(y,x,q,p))$	$v_y(\mathsf{CC}(x,y,p,q)) + v_y(\mathsf{CC}(y,x,q,p))$	$v_y(\mathsf{CC}(x,y,p,q)) \geq 0$
complement	$v_y(\mathsf{CC}(x,y,p,q)) + v_y(\mathsf{CC}(y,x,q,p))$	$v_y(\mathsf{C}(x,y,q)) + v_y(\mathsf{CC}(y,x,q,p))$	$v_y(\mathsf{C}(x,y,q))$ $-v_y(\mathsf{CC}(x,y,p,q)) \geq 0$
accept	$v_y(\mathsf{CC}(y,x,q,p))$	$v_y(\mathsf{CC}(y,x,q,p)) + v_y(\mathsf{C}(x,y,q))$	$v_y(\mathsf{C}(x,y,q)) \geq 0$
ponens	$v_y(\mathsf{C}(y,x,p)) + v_y(\mathsf{CC}(x,y,p,q))$	$v_y(\mathsf{C}(y,x,p)) + v_y(\mathsf{C}(x,y,q))$	$v_y(\mathsf{C}(x,y,q))$ $-v_y(\mathsf{CC}(x,y,p,q)) \geq 0$
discharge-C	$v_y(\mathsf{C}(y,x,p)) + v_y(\mathsf{C}(x,y,q))$	$v_y(\mathsf{C}(y,x,p)) + v_y(q)$	$v_y(q) - v_y(\mathsf{C}(x,y,q)) \geq 0$

running example, the customer may say that it is willing to pay if the merchant delivers the goods. If the merchant commits to delivering the goods, the customer can commit to paying.

We can see exactly how the challenge rule is rational for x. If x knows itself to be honest, $v_x(q) = v_x(\mathsf{C}(x,y,q))$, but if x does not fully trust y, $v_x(\mathsf{C}(y,x,p)) < v_x(p)$. Thus an agent who does not trust another would find it rational to challenge. Now if each agent does not trust the other, there would be a deadlock. The risk tolerance of an agent corresponds to how much it is willing to take on. Given factors such as the valuation assigned to reputable behavior, if the difference in valuations is within the agents' risk tolerance, the agents can still make progress.

4.2 Applying the Concession Rules

Figure 2 depicts the evolution of a protocol based on the concession rules described before. To understand the effect of the rules, we consider the values of the states that come about through the application of these rules. Table 2 tabulates the valuation changes for x induced by each rule.

EXAMPLE 6. *Consider a valuation function v_C for Customer in the purchasing protocol. Customer's goal is to receive goods, hence G(Customer, goods) holds. Customer's valuation of its benefits can be shown with the following $v_C(goods) = 2$, $v_C(C(M,C,goods)) = 1$, $v_C(CC(M,C,pay,goods)) = 0.5$. The merchant will expect the customer to pay in exchange for goods. Customer's valuation of its risks is as follows: $v_C(pay) = -1$; $v_C(C(C,M,pay)) = -0.5$; and $v_C(CC(C,M,goods,pay)) = -0.25$. In other words, the customer values receiving goods more than making the payment. v_C is a coherent valuation function since it satisfies the conditions of Definition 4.*

Let's consider the values of some states.

- *Consider the state where the customer has made the conditional commitment $\mathsf{CC}(C,M,goods,pay)$ and the merchant made a commitment to deliver (accept), $\mathsf{C}(M,C,goods)$. The v_C for this state is $v_C(\mathsf{CC}(C,M,goods,pay))+$ $v_C(\mathsf{C}(M,C,goods)) = 0.75$.*
- *Consider the state where the customer and the merchant are committed to each other to sending the goods and the payment, respectively. Thus, $\mathsf{C}(C,M,pay)$ and $\mathsf{C}(M,C,goods)$ hold. The v_C for this state is $v_C(\mathsf{C}(C,M,pay))+$ $v_C((M,C,goods)) = 0.5$. This is, in general, a less desirable state for the customer since now the customer is committed to paying, whereas in the previous case, the customer did not have such a commitment.*
- *Consider the state where the customer is committed to paying and the merchant has delivered the goods. In general, this is a better state than the above states since the customer has reached its goal of receiving the goods with the cost of committing to pay. The valuation of this state is $v_C(\mathsf{C}(C,M,pay))+ v_C(goods) = 1.5$.* ▮

Table 5: Social welfare of states for the purchasing protocol

State	v_C	v_M	(s)
CC(C, M, goods, pay)	-0.25	0.50	0.25
CC(C, M, goods, pay) ∧ CC(M, C, pay, goods)	0.25	0.25	0.50
CC(C, M, goods, pay) ∧ C(M, C, goods)	0.75	0	0.75
C(C, M, pay) ∧ C(M, C, goods)	0.50	0.50	1.00
C(C, M, pay) ∧ goods	1.50	0	1.50
pay ∧ goods	1.00	1.00	2.00

EXAMPLE 7. *Table 5 shows the increasing social welfare for different states in Figure 2, using the above valuation function for the customer and the converse function for the merchant.* ▮

4.3 Properties of Commitment Concession

The concession rules enable agents to make slow steps toward their goals. Some of the rules create transitions that decrease the valuation of agents, hence, they are truly *concession* rules.

EXAMPLE 8. *Consider accept in Table 2. Before x executes accept, $\mathsf{CC}(y,x,q,p)$ holds, meaning that x will receive p in return*

Table 4: Social welfare as induced by various concession rules

Rule	w **before**	w **after**
create-CC	0	$v_x(\mathsf{CC}(x,y,p,q)) + v_y(\mathsf{CC}(x,y,p,q))$
challenge	$v_x(\mathsf{CC}(x,y,p,q)) + v_y(\mathsf{CC}(x,y,p,q))$	$v_x(\mathsf{CC}(x,y,p,q)) + v_x(\mathsf{CC}(y,x,q,p))$ $+v_y(\mathsf{CC}(x,y,p,q)) + v_y(\mathsf{CC}(y,x,q,p))$
complement	$v_x(\mathsf{CC}(x,y,p,q)) + v_x(\mathsf{CC}(y,x,q,p))$ $+v_y(\mathsf{CC}(x,y,p,q)) + v_y(\mathsf{CC}(y,x,q,p))$	$v_x(\mathsf{C}(x,y,q)) + v_x(\mathsf{CC}(y,x,q,p))$ $+v_y(\mathsf{C}(x,y,q)) + v_y(\mathsf{CC}(y,x,q,p))$
accept	$v_x(\mathsf{CC}(x,y,p,q)) + v_y(\mathsf{CC}(x,y,p,q))$	$v_x(\mathsf{CC}(x,y,p,q)) + v_x(\mathsf{C}(y,x,p))$ $+v_y(\mathsf{CC}(x,y,p,q)) + v_y(\mathsf{C}(y,x,p))$
ponens	$v_x(\mathsf{CC}(y,x,q,p)) + v_x(\mathsf{C}(x,y,q))$ $+v_y(\mathsf{CC}(y,x,q,p)) + v_y(\mathsf{C}(x,y,q))$	$v_x(\mathsf{C}(y,x,p)) + v_x(\mathsf{C}(x,y,q))$ $+v_y(\mathsf{C}(y,x,p)) + v_y(\mathsf{C}(x,y,q))$
discharge-C	$v_x(\mathsf{C}(y,x,p)) + v_x(\mathsf{C}(x,y,q))$ $+v_y(\mathsf{C}(y,x,p)) + v_y(\mathsf{C}(x,y,q))$	$v_x(\mathsf{C}(y,x,p)) + v_x(q)$ $+v_y(\mathsf{C}(y,x,p)) + v_y(q)$

of q. However, by execution of this rule, additionally it becomes committed to q ($\mathsf{C}(x,y,q)$). Since, $v_x(\mathsf{C}(x,y,q) \leq 0$, x has moved to a state with lower valuation. ▮

Even though there is a decrease in the valuations for the acting agent, the decrease is small. An agent would take the risk if the amount of risk is within some risk tolerance.

LEMMA 1. *Let x be an agent, s be a state where agent x executes one of create-C, challenge, ponens, accept, complement, and discharge-C, moving the protocol state to s'. Then, $v_x(s) > v_x(s')$.*
Proof. *Follows from Table 2.*

LEMMA 2. *Let x and y be agents, s be a state where agent x executes one of create-C, challenge, ponens, accept, complement, and discharge-C, moving the protocol state to s'. Then, $v_y(s) < v_y(s')$.*
Proof. *Follows from Table 3.*

Lemma 1 and Lemma 2 jointly show that when an agent performs one of the concession moves, it decreases its valuation, but increases the other party's valuation. In principle, it would be irrational for the agents to make transitions that lower their valuation. However, by making small devaluations at each step increases the social welfare of the system, thus guaranteeing that agents reach a final state. We next show that if all the agents choose to execute one of these rules at every move of the protocol, then the protocol will reach a final state, where all the goal propositions are satisfied.

To show this, we need a metric to measure the progress of transitions in a protocol. In connection with commitments, the natural measure need not be a distance metric, because it need not be symmetric. For this reason, a more general notion, that of a *quasidistance* is needed.

Table 6: Inductive definition of quasidistance between states based on social welfare

From	To	Quasidistance (Q)
$\{\ \}$	$\{p\}$	$w(p)$
$\{\ \}$	$\{c\}$	$w(c)$
$\{\ \}$	$\{cc\}$	$w(cc)$
$\{q\}$	$\{p\}$	$w(p) - w(q)$
$\{c\}$	$\{p\}$	$w(p) - w(c)$
$\{cc\}$	$\{p\}$	$w(p) - w(cc)$
F_1	$T_1 \cup T_2$	$Q(F_1, T_1) + Q(F_1, T_2)$
$F_1 \cup F_2$	T_1	$\min(Q(F_1, T_1), Q(F_2, T_1)$

DEFINITION 7. *Table 6 gives the inductive rules to calculate the quasidistance Q based on social welfare. $Q(s, s')$ shows the quasidistance from state s to state s' in a protocol. In this table, $c = \mathsf{C}(y,x,p)$ refers to a base-level commitment; $cc = \mathsf{CC}(x,y,q,p)$ to a conditional commitment; and p and q are atomic propositions (which could be the same). F_1, T_1, and T_2 represent a set of propositions or commitments. The quasidistance from a set F_1 to the union of two sets T_1 and T_2 ($Q(F_1, T_1 \cup T_2)$) is the summation of the quasidistances from T_1 to F_1 and from T_1 to F_2, since both T_1 and T_2 need to be achieved. However, the quasidistance from a union of two sets F_1 and F_2 to a set T_1 ($Q(F_1 \cup F_2, T_1)$) is the minimum of the quasidistances between F_1 and T_1 and F_2 and T_1, because achieving T_1 from one of F_1 or F_2 is enough.*

As explained in Section 2, in principle, a protocol can end at any state where there do not exist any base-level commitments. However, a preferred final state among such states is one where all enacting agents have achieved their goals. The following definition captures this intuition.

DEFINITION 8. *Let x and y be two agents that enact a protocol. State s is a* final state *if both agents receive their credits at s.*

LEMMA 3. *Let s be a final state and x and y be two agents enacting the protocol. Then, $w(s) > 0$.*
Proof. *Based on our assumption of rationality, each agent enacts the protocol if the final state is beneficial for itself. Since, both $v_x(s) > 0$ and $v_y(s) > 0$ hold, $w(s) > 0$.*

LEMMA 4. *Let s and s' be two states in the protocol, such that s is a final state. If $Q(s, s') > 0$, then s' is a final state as well.*
Proof. *If s''s quasidistance to s is positive, then s' has higher social welfare. This means that all the parties have received more than their expected credits. By Definition 8, s' is a final state, too.*

LEMMA 5. *Let s and s' be two consecutive states such that s' results when an agent applies one of create-CC, challenge, ponens, accept, complement, and discharge-C. Then, $Q(s, s') > 0$.*
Proof. *Follows from Table 4. With each move, each agent decreases its valuation but increases the other party's valuation. Recall that the valuations of the agents are not symmetric (i.e., $v_x(p) + v_y(p) > 0$). In the cases shown, a decrease in one agent's valuation, increases the other's valuation more, increasing the total. Thus, the social welfare increases.*

THEOREM 6. *A commitment protocol converges to a final state if every agent exercises one of create-CC, challenge, ponens, accept, complement, and discharge-C at every move.*

Proof. *By Lemma 3, we know that the final state has a positive social welfare. By Lemma 5, applying* create-CC, challenge, ponens, accept, complement, *and* discharge-C *will increase the social welfare repeatedly. The protocol will eventually reach a state where social welfare is greater than or equal to all agents receiving their credits. By Lemma 4, this is a final state.*

5. DISCUSSION

Concession moves as described here can be implemented (1) in a separate module independently of the domain protocol; (2) as a separate interaction protocol; or (3) embedded into a domain protocol. The last is a particularly interesting situation. We are given a protocol that serves some domain purpose. This protocol can be enacted by a variety of agents in a variety of circumstances. For example, a purchase protocol among agents who trust each other or function within a trust environment may proceed in a simple manner sending each other the relevant domain messages. By contrast, agents who enact the same protocol in different circumstances might well have to carry out negotiations about the key messages of the domain protocol. Such negotiations can be generated via the concession approach described in this paper.

Wan and Singh study commitments in reaching multiparty agreements [10]. They develop agreement derivation rules that transform a multiparty agreement into actions that need to be carried out by individual agents. They further propose an algorithm for detecting and resolving deadlocks that may arise when the actions are applied arbitrarily. Their resolution is based on an agent's transforming its conditional commitment to an unconditional commitment based on trust in the other party. This is similar to the intuition behind the accept rule proposed in this paper.

Karunatillake *et al.* propose an approach for argumentation-based negotiation based on social commitments [5]. The agents that negotiate are bound to organizational roles and are influenced by social relationships. The approach enumerates possible rules that can be applied when conflicts occur. These rules are used to reject or accept proposals as well as to enforce the social relations. However, the rules in Karunatillake *et al.*'s approach are not devised for negotiating the content of the commitments in small increments as we have done here.

McBurney and Parsons handle e-commerce transactions using posit spaces protocol that consists of propose, accept, delete, suggest_revoke, and ratify_revoke [6]. The usage of propose and accept locution resembles the conditional commitments in commitment protocols. The delete locution corresponds to the release, or discharge operation. Suggest_revoke and ratify_revoke enable canceling of posits. These locutions are useful for negotiation, but are not sufficient to enable negotiation with concessions. For example, commitments can be accepted or rejected by these offers, but the content of a commitment cannot be restricted with these locutions.

Flores and Kremer develop a commitment-based approach for designing conversation protocols [4]. The design phase includes steps for identifying agent roles, agent actions, and so on. Our focus in this paper is complementary to that Flores and Kremer since we design concession rules that will enable agents to reach their final states by taking small risks at a time.

Alberti *et al.* specify interaction protocols using social integrity constraints [1]. Given a set of event occurrences, each agent computes a set of expected events based on the social integrity constraints. Social integrity constraints help agents reach a common goal. However, contrary to our proposal of negotiating with commitments at runtime, social integrity constraints are static throughout a protocol enactment.

Endriss studies the properties of monotonic concession protocols in multilateral negotiations from a game-theoretic point of view [3]. Endriss proposes different definitions for multilateral concession and analyzes protocols in terms of termination, liveness, and so on. It would be interesting to study how his definitions of concession would apply to commitment protocols.

This paper has sought to combine considerations of rationality with those of commitment protocols. Many interesting research challenges arise from this combination, including the treatment of more general kinds of concession. Further, valuation functions with different characteristics (e.g., superadditive or subadditive) can be considered. It would also be helpful to extend the approach of this paper to arbitrarily nested commitments.

6. REFERENCES

[1] M. Alberti, M. Gavanelli, E. Lamma, P. Mello, and P. Torroni. Modeling interactions using social integrity constraints: A resource sharing case study. In *Proc. Intl. Workshop on Declarative Agent Languages and Technologies (DALT), LNAI 2990*, pages 243–262. Springer-Verlag, 2004.

[2] C. Castelfranchi. Commitments: From individual intentions to groups and organizations. In *Proc. AAAI-93 Workshop on AI and Theories of Groups and Organizations: Conceptual and Empirical Research*, 1993.

[3] U. Endriss. Monotonic concession protocols for multilateral negotiation. In *Proc. 5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-06)*, pages 392–399. ACM Press, May 2006.

[4] R. A. Flores and R. C. Kremer. A principled modular approach to construct flexible conversation protocols. In *Proc. Canadian Conference on AI, LNCS 3060*, pages 1–15. Springer-Verlag, 2004.

[5] N. C. Karunatillake, N. R. Jennings, I. Rahwan, and T. J. Norman. Argument-based negotiation in a social context. In *Proc. AAMAS Workshop on Argumentation*, pages 74–88, May 2005.

[6] P. McBurney and S. Parsons. Posit spaces: a performative model of e-commerce. In *Proc. the 2nd Intl. Joint Conf. Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 624–631. ACM Press, 2003.

[7] P. Pasquier and B. Chaib-draa. Modelling the links between social commitments and individual intentions. In *Proc. 5th Intl. Joint Conf. Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1372–1374. ACM Press, May 2006.

[8] J. S. Rosenschein and G. Zlotkin. *Rules of Encounter*. MIT Press, Cambridge, MA, 1994.

[9] M. P. Singh. An ontology for commitments in multiagent systems: Toward a unification of normative concepts. *Artificial Intelligence and Law*, 7:97–113, 1999.

[10] F. Wan and M. P. Singh. Formalizing and achieving multiparty agreements via commitments. In *Proc. of the 4th Intl. Joint Conf, on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 770–777. ACM Press, 2005.

[11] M. Winikoff, W. Liu, and J. Harland. Enhancing commitment machines. In *Proc. 2nd Intl. Workshop Declarative Agent Languages and Technologies (DALT), LNAI 3476*, pages 198–220. Springer-Verlag, 2005.

[12] P. Yolum and M. P. Singh. Flexible protocol specification and execution: Applying event calculus planning using commitments. In *Proc. 1st Intl. Joint Conf. Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 527–534. ACM Press, 2002.