A Protocol for Arguing about Rejections in Negotiation

Jelle van Veenen[1] and Henry Prakken[23]

[1] Faculty of Law, Tilburg University, The Netherlands,
j.vanveenen@uvt.nl
[2] Institute of Information and Computing Sciences, Utrecht University, The Netherlands,
henry@cs.uu.nl
[3] Centre for Law & ICT, Faculty of Law, University of Groningen, The Netherlands

Abstract One form of argument-based negotiation is when agents argue about why an offer was rejected. If an agent can state a reason for a rejection of an offer, the negotiation process may become more efficient since the other agent can take this reason into account when making new offers. Also, if a reason for rejection can be disputed, the negotiation process may be of higher quality since flawed reasons may be revised as a result. This paper presents a formal protocol for negotiation in which reasons can be asked and given for rejections and in which agents can try to persuade each other that a reason is or is not acceptable. The protocol is modelled as a persuasion dialogue game embedded in a negotiation protocol. It has a social semantics since the protocol does not refer to the internal state of negotiating agents.

1 Introduction

Recently argumentation-based approaches to negotiation have become popular (see [1] for an overview and motivation). The idea is that if negotiating agents exchange reasons for their proposals and rejections, the negotiation process may become more efficient and the negotiation outcome may be of higher quality. This paper especially focuses on reasons given for rejections of proposals. If an agent explains why he rejects a proposal, the other agent knows which of her future proposals will certainly be rejected so she will not waste effort at such proposals. Thus efficiency is promoted. In such exchanges, reasons are not only exchanged, they can also become the subject of debate. Suppose a car seller offers a Peugeot to the customer but the customer rejects the offer on the grounds that French cars are not safe enough. The car seller might then try to persuade the customer that he is mistaken about the safety of French cars. If she succeeds in persuading the customer that he was wrong, she can still offer her Peugeot. Thus the quality of the negotiation is promoted, since the buyer has revised his preferences to bring them in agreement with reality.

This example illustrates that a negotiation dialogue (where the aim is to reach a deal) sometimes contains an embedded persuasion dialogue (where the aim is to resolve a conflict of opinion). The aim of this paper is to formulate a protocol for negotiation with embedded persuasion dialogues about the reasons for rejecting a proposal. The key idea is that the propositional commitments incurred by the agents in the embedded persuasion dialogue constrain their behaviour in the surrounding negotiation dialogue. In agreement with the current trend [2] we intend the combined protocol to have a social semantics. For this reason we will completely abstract from the internal design of the communicating agents; in particular, the protocol will only refer to the agents' publicly observable behaviour within a dialogue. According to [2] a social semantics is desirable for agent interaction protocols since if a protocol refers to an agent's mental state, there is no gurantee that an outside observer can verify whether the agent complies with a protocl.

The main novelty of the present research lies in the fact that current protocols for argument-based negotiation only allow arguments supporting proposals. One exception is [3], which also allows arguments about rejections. However, their protocol does not have a social semantics, since whether an agent is allowed to assert a claim or an argument partly depends on their internal mental state. The protocol has some other limitations, which will be discussed in Section 5.

Our proposal will be stated in a dialogue game form. It will combine a negotiation protocol and language of [4] with a persuasion protocol based on [5], which adapts and extends work of [6]. In the following sections we will first introduce these two systems and sketch the underlying argumentation logic that we will use. Then we will present our combined protocol, investigate some of its formal properties and illustrate it with an example.

2 The building blocks

In this section we present the negotiation and persuasion system that we aim to combine. Both systems are formulated as a dialogue game. Dialogue games formulate principles for coherent dialogue, and coherence depends on the goal of a dialogue. The goal of negotiation dialogues is to reach agreement on the division of scarce resources [1] and the goal of persuasion dialogues is to resolve a conflict of opinion [7]. Formal dialogue games have a *topic language* L_t with a logic $\mathcal{L}$, and a *communication language* L_c with a *protocol* P. The protocol specifies the allowed moves at each point in a dialogue. A dialogue system also has *effect rules*, which specify the effects of utterances on the participants' commitments, and *outcome rules*, defining the outcome of a dialogue.

2.1 A language and protocol for multi-attribute negotiation

The negotiation system we will use is that of Wooldridge and Parsons [4]. The **negotiation topic language** L_t^n of this system assumes that in a negotiation agents try to reach agreement over the values of a finite set $V = \{v_1, ..., v_m\}$ of *negotiation issues*. Each issue v can be assigned at most one value from a range $C(v)$ of values. An *outcome* of a negotiation is an assignment of values to a subset of V. A proposal is expressed in a subset of the language of first-order logic as a conjunction of expressions of the form vRc, where $v \in V$ and $c \in C(v)$ or $c = ?$, (where ? technically is a free variable, capturing that the issue has not been assigned a value) and R denotes one of the relations $=, <, >, \leq$ or $\geq$.

The **negotiation communication language** L_c^n can be used to talk about proposals. The left column of table 1 shows the speech acts that agents can perform and the right column their possible replies. The formulas φ and φ' are elements of L_t^n. *Request(φ)* is a request for an offer. Here φ typically is wholly or partially uninstantiated (i.e., it may contains occurrences of ?): the speech act $request(price = ? \wedge warranty = 12)$ can be read as "What is the price if I want a 12 months warranty?". The speech act *offer*(φ) makes a fully instantiated proposal φ, and with *accept*(φ) an agent accepts an offer φ made by another agent. With *reject*(φ) such an offer is rejected. With *withdraw* an agent withdraws from the negotiation.

We next outline the **negotiation protocol** of [4] for this language, with notation slightly adapted to our purposes. A negotiation takes place between two agents, one of whom starts with either an *offer* or a *request*. The agents then take turns after each utterance, selecting their replies from Table 1. As the table indicates, a negotiation terminates when an agent accepts an offer or withdraws from the negotiation. Finally, moves may not be repeated by the same player.

Table 1. Speech acts and replies in L_c^n

Acts	Replies:
request(φ)	*offer(φ')*
offer(φ)	*offer(φ')* or *accept(φ)* or *reject(φ)* or *withdraw*
reject(φ)	*offer(φ')* or *withdraw*
accept(φ)	end of negotiation
withdraw	end of negotiation

($\varphi \neq \varphi'$)

To ensure that the offers exchanged during a negotiation and its outcome are related to an initial request, we add the following rule to the protocol of [4]:

– If *request*(φ) is the initial request of a dialogue then for any move *offer*(ψ) in the dialogue:
  - ψ is logically consistent with φ; and
  - ψ contains at least the same issues as φ.

Since issues have at most one value, this rule implies that an instantiated part of a request cannot be changed by an offer (but the offer may contain more issues than the request). Therefore:

Proposition 1. *If a negotiation that starts with a request terminates with acceptance of an offer, that offer is consistent with and fully instantiates the request.*

We illustrate the system with an example in which two agents, Paul (P) and Olga (O), negotiate over the sale of a car. The dialogue starts when Paul requests to buy a car, and shows that he is interested in the brand and the price.

P_1: $request(brand = ? \wedge price = ?)$
O_2: $offer(brand = peugeot \wedge price = 10000)$
P_3: $reject\ (brand = peugeot \wedge price = 10000)$
(Olga has offered a Peugeot for 10000, but Paul has rejected the offer. Olga makes him another offer.)
O_4: $offer(brand = renault \wedge price = 8000 \wedge stereo{=}\ yes)$
P_5: $reject(brand = renault \wedge price = 8000 \wedge stereo{=}\ yes)$
O_6: $offer(brand = audi \wedge price = 10000)$
P_7: $accept(brand = audi \wedge price = 10000)$
(Olga offers a Renault with stereo for 8000. Paul again rejects after which Olga offers a non-French car for 10000. Paul accepts and the dialogue terminates. Move O_4 illustrates that an offer may introduce additional issues, for instance, to make an offer more attractive or to make a trade-off possible.)

2.2 The underlying argumentation logic

We next present the logical elements assumed by our persuasion protocol, i.e., the persuasion topic language L_t^p and its logic $\mathcal{L}$. In doing so we will abstract from details of the language and inference rules wherever possible, to allow for different instantiations of the logic and language. Thus we in fact specify a set of constraints on L_t^p and $\mathcal{L}$ assumed by our persuasion dialogue system.

Much other work on argument-based dialogue, e.g. [3,8], regards arguments as classical proofs from consistent sets of propositional formulas and allows classical inconsistency of the premises of two arguments as the only source of attack on arguments. We argue that the present application requires a richer language and notions of argument and attack. The topic language L_t^p must include a suitable subset of first-order predicate logic, to express arguments about values for negotiation issues. Since in persuasion dialogues arguments are often attacked by counterarguments, the logic $\mathcal{L}$ must be a logic for defeasible argumentation, or 'argumentation system' for short (cf. [9]). We want our system to be an instance of the well-understood abstract framework of [10], in particular of his grounded semantics (also used by e.g. [11,12]), since this semantics can be easily incorporated into a persuasion dialogue game. Since the arguments exchanged in persuasion dialogues are often constructed stepwise during a dialogue in reply to challenges of the premises, the argumentation system must allow for a tree structure of arguments, where inference rules are chained into trees. As for notation, $prem(A)$ and $conc(A)$ denote the premises and conclusion of argument A, i.e. the leaves and root of the tree structure. Furthermore, since arguments exchanged in persuasion dialogues are often based on defeasible argumentation schemes (such as schemes for practical reasoning or default reasoning), the rules for constructing arguments must include defeasible as well as deductive inference rules. Each defeasible inference rule comes with one or more *undercutters*, which specify the circumstances under which the inference rule cannot be applied. Accordingly, a defeasible argument can be defeated in two ways. It can be *rebut* with an argument for the opposite conclusion, while it can be *undercut* with an argument why an inference rule does not apply in the given circumstances. To be successful, an attack should be of a certain strength. In the present paper, we will not discuss issues of strength and therefore implicitly assume a given measure of relative strength between arguments. Also, since our persuasion dialogue system is intended for any underlying logic satisfying the above constraints, we will not further specify the defeasible inference rules here but rather introduce them semiformally when discussing our examples. For technical details the reader is referred to e.g. [13] and [9].

Given a set of arguments and a binary defeat relation defined over it, an argumentation system classifies the arguments into justified, defensible and overruled arguments. Our persuasion system presupposes a game-theoretic formulation of Dung's grounded semantics [14,11]. The proponent and opponent of a certain argument play a game where proponent starts with an argument he wants to defend and then both players take turns, defeating the preceding argument with a counterargument. A player wins if the other player has run out of moves. Now an argument is justified if proponent has a winning strategy in a game starting with the argument; and a proposition is justified if it is the conclusion of a justified argument. This game can be optimised in several ways (see e.g. [11]) but in order to focus on the essence we leave them undiscussed here.

2.3 A dialogue game for persuasion

We now present a dialogue game for persuasion. As noted above, this game is an instance of the framework of [5], which adapts and further develops the system of [6]. We are particularly interested in using this framework's idea of reply structure on the communication language and its notions of dialogical status of relevance. A crucial feature of our game is that its protocol is flexible in that it allows for alternative replies to moves and for postponement of replies, sometimes even indefinitely. This is important since when an agent sees that a line of attack or defence fails, s/he should be allowed to play other available lines of attack or defence. However, in order to still ensure a strong focus of dialogues this flexibility is constrained by the notion of relevance, to be defined below.

The dialogue game will be presented here in detail since its format plays a crucial role in Section 3 in the combination of the persuasion and negotiation dialogue game. Dialogues are between a *proponent* P and *opponent* O of a single *dialogue topic* $t \in L_t^p$. The game is based on the following ideas. Each dialogue move except the initial one replies to one earlier move in the dialogue of the other party (its *target*). Thus a dialogue can be regarded in two ways: as a sequence (reflecting the order in which the moves are made) and as a tree (reflecting the reply relations between the moves). Each replying move is either an *attacker* or a *surrender*. For instance, a $claim(p)$ move can be attacked with a $why(p)$ move and surrendered with a $concede(p)$ move. And a $why(p)$ move can be attacked with an $argue(A)$ move where A is an argument with conclusion p, and surrendered with a $retract(p)$ move. When s is a surrendering and s' is an attacking reply to s'', we say that s' is an *attacking counterpart* of s. The **persuasion communication language** L_c^p is specified in Table 2. In this

Table 2. Speech acts and replies in L_c^p.

Acts	Attacks	Surrenders
$claim(\varphi)$	$why(\varphi)$	$concede(\varphi)$
$why(\varphi)$	$argue(A)$ $(conc(A) = \varphi)$	$retract(\varphi)$
$argue(A)$	$why(\varphi)$ $(\varphi \in prem(A))$ $argue(B)$ (B defeats A)	$concede(\varphi)$ $(\varphi \in prem(A)$ or $\varphi = conc(A))$
$concede(\varphi)$		
$retract(\varphi)$		

table, φ is from L_t^p and arguments A and B are well-formed arguments from $\mathcal{L}$, while defeat relations between arguments are determined according to $\mathcal{L}$. Thus the proof theory of $\mathcal{L}$ is embedded in the persuasion protocol.

The protocol for L_c^p is defined in terms of the notion of a **dialogue**, which in turn is defined with the notion of a **move**.

Definition 1.

- *The set M of moves is defined as $\mathbb{N} \times \{P, O\} \times L_c^p \times \mathbb{N}$, where the four elements of a move m are denoted by, respectively:*
  - *$id(m)$, the identifier of the move,*
  - *$pl(m)$, the player of the move,*
  - *$s(m)$, the speech act performed in the move,*
  - *$t(m)$, the target of the move.*
- *The set of dialogues, denoted by $M^{\leq\infty}$, is the set of all sequences $m_1, \ldots, m_i, \ldots$ from M such that*
  - *each i^{th} element in the sequence has identifier i,*
  - *$t(m_1) = 0$;*
  - *for all $i > 1$ it holds that $t(m_i) = j$ for some m_j preceding m_i in the sequence.*

  The set of finite dialogues, denoted by $M^{<\infty}$, is the set of all finite sequences that satisfy these conditions. For any dialogue $d = m_1, \ldots, m_n, \ldots$, the sequence $m_1, \ldots, m_i$ is denoted by d_i, where d_0 denotes the empty dialogue.

When $t(m) = id(m')$ we say that m replies to m' in d and that m' is the target of m in d. We sometimes slightly abuse notation and let $t(m)$ denote a move instead of just its identifier. When $s(m)$ is an attacking (surrendering) reply to $s(m')$ we also say that m is an attacking (surrendering) reply to m'.

The semantics for L_c^p is defined in axiomatic style as a set of precondition-postcondition rules. In fact, as we will see below, the only precondition for each move is that it is legal at this point in the dialogue according to the **protocol**. The protocol is stated in a form defined by [15]:

Definition 2. *(Protocols for games.) A protocol on M is a function P with domain a nonempty subset of $M^{<\infty}$ taking subsets of M as values. The elements of $dom(P)$ (the domain of P) are called the legal finite dialogues. The elements of $P(d)$ are called the moves allowed after d. If d is a legal dialogue and $P(d) = \emptyset$, then d is said to be a terminated dialogue. P must satisfy the following condition: for all finite dialogues d and moves m, $d \in dom(P)$ and $m \in P(d)$ iff $d, m \in dom(P)$.*

Definition 3. *(The protocol P^p for L_c^p.) For all moves m it holds that $m \in P^p(d)$ if and only if m satisfies all of the following rules:*

- R_1: *$pl(m) = T(d)$;*[4]
- R_2: *If $d \neq d_0$ and $m \neq m_1$, then $s(m)$ is a reply to $s(t(m))$ according to L_c^p;*
- R_3: *If m replies to m', then $pl(m) \neq pl(m')$;*
- R_4: *If there is an m' in d such that $t(m) = t(m')$ then $s(m) \neq s(m')$;*
- R_5: *If $d = d_0$, then $s(m)$ is of the form $claim(\varphi)$;*
- R_6: *If $s(m) = retract(\varphi)$, then $C_s(d,m) \nvdash \varphi$;*
- R_7: *$C_s(d,m)$ is consistent;*
- R_8: *if m is a replying move, then m is relevant in d.*

(for relevance see further below). Further rules could be added, for instance, to prevent circular dialogues [7,16], but to focus on the essence we will leave such rules undiscussed here.

R_1 says that the player of a move must be to move. R_2-R_4 formalise the idea of a dialogue as a move-reply structure that allows for alternative replies. R_5 says that each dialogue begins with a claim; the initial claim is the *topic* of the dialogue. R_6 requires retractions to be successful and R_7 requires the players to keep their commitments consistent. Finally, rule R_8 says that each replying move must be relevant. This crucial element of the protocol requires some explanation.

Relevance is defined in terms of the *dialogical status* of a move, which in turn is recursively defined in terms of the nature of its replies. A move is *in* iff it is surrendered or else if all its attacking replies are out. (This implies that a move without replies is in). And a move is *out* if it has an attacking reply that is in. With this concept of dialogical status a notion of relevance can be defined. A move is *relevant* if it replies to a relevant target. And a move is a *relevant target* if making it out changes the dialogical status of the initial move of the dialogue. Together with Definition 3 these definitions imply that a move is a relevant target for proponent (opponent) if making it out makes the initial move

[4] $T(d)$ denotes the player whose turn it is to move in d.

in (out). Accordingly we say that P *currently wins* d if m_1 is in and O *currently wins* if m_1 is out.

Figure 1 (with only attacking replies) illustrates the notion of relevance. A

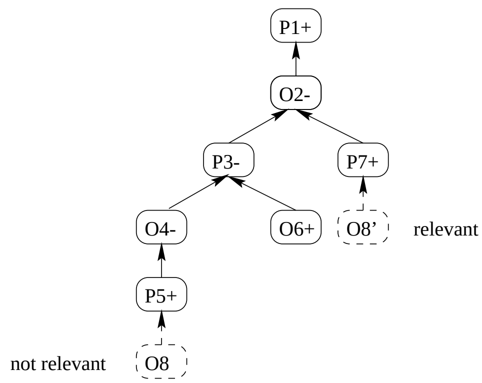

Figure 1. Relevance of moves.

move labelled + is in and a move labelled − is out. P_5 is not a relevant target for O: although making P_5 out makes O_4 in, P_3 was already out because of O_6 and therefore O_2 stays out because of P_7, so that P_1 stays in. However, P_7 is a relevant target for O: making P_7 out makes O_2 in since its only attacking reply is now out; then P_1 is out since it now has an attacking reply that is in.

The requirement of relevance comes with a **turntaking rule** T that the turn switches as soon as a player has changed the dialogical status of the initial move (below p is a variable ranging over $\{P, O\}$ and $\overline{p}$ denotes O if $p = P$ and P if $p = O$). Formally, T is a function

- $T : M^{<\infty} \longrightarrow \{P, O\}$

such that $T(d_0) = P$ and if $d \neq d_0$ then $T(d) = p$ iff $\overline{p}$ currently wins d.

The rationale of this rule is that as soon as a player has changed the dialogical status of m_1, he has no relevant moves any more so to avoid premature termination the turn should shift to the other party.

The **effect rules** are defined as a function of the following type:

- C: $\{P, O\} \times M^{<\infty} \longrightarrow \mathcal{P}(L_t^p)$.

$C_p(d)$ denotes the commitments of player p in the dialogue d. The following commitment rules for L_c^p seem uncontroversial and can be found throughout the literature. (Below s denotes the speaker of the move; effects on the other parties' commitments are only specified when a change is effected; finally, d, m stands for the dialogue starting with dialogue d and continuing with move m.)

- If $s(m) = claim(\varphi)$ then $C_s(d, m) = C_s(d) \cup \{\varphi\}$

– If $s(m) = why(\varphi)$ then $C_s(d, m) = C_s(d)$
– If $s(m) = concede(\varphi)$ then $C_s(d, m) = C_s(d) \cup \{\varphi\}$
– If $s(m) = retract(\varphi)$ then $C_s(d, m) = C_s(d) - \{\varphi\}$
– If $s(m) = argue(A)$ then $C_s(d, m) = C_s(d) \cup prem(A) \cup \{conc(A)\}$

The **axiomatic semantics** of the system then is as follows: for each move m and dialogue d:

precondition: $m \in P^p(d)$
postcondition: as specified by $C_p(d, m)$.

To give a feel for how dialogues evolve in this system, we now list a few properties of the system (see [5] for more details). Firstly, a *turn* of a player always consists of zero or more surrenders followed by a single attack. Further, the turn shifts to the opponent if the initial move is made in while it shifts to the proponent if the initial move is made out. It also follows that a dialogue terminates only if the status of the initial move is *against* the player to move (out for the proponent and in for the opponent). So if a dialogue terminates when player p is to move, p can be said to have *lost* the dialogue. Moreover, it can be shown that a dialogue terminates if and only if either proponent has surrendered to opponent's first move by retracting the dialogue topic or opponent has surrendered to proponent's first move by conceding the dialogue topic. Finally, a 'fairness' and 'soundness' result can be proven about the relation between the dialogical status of the initial move on the one hand and the underlying logic on the other. Under certain conditions the initial move is in just in case the initial claim is defeasibly implied by the 'defended' arguments exchanged by the parties, that is, by the arguments without challenged premises.

Dialogues are not guaranteed to terminate, since the opponent can always continue challenging the proponent's premises. This is a consequence of the fact that the protocol ignores the agents' internal design so that their knowledge bases are not guaranteed to remain fixed during a dialogue. In our opinion this is not a bug but a feature of the protocol since in many realistic settings the agents' knowledge changes during a dialogue. For instance, they may ask advice of third parties, consult databases or make new observations.

3 Negotiation and argumentation

In the previous sections we introduced protocols for negotiation and persuasion. We now combine them in a way that allows persuasion dialogues to be embedded in negotiation dialogues. In a negotiation dialogue it is the reject move that shows that there is a conflict between the preferences of an agent and the offer that it receives. By starting a persuasion dialogue, the offerer can question the reasons that the offeree has for rejecting. Statements made during persuasion invoke commitments that reflect the preferences of the agents. These commitments are used to restrict further negotiations.

In formally realising the combination of the two dialogue systems, the key idea is to reformulate the negotiation system in the format of Section 2.3 so that the mechanisms of relevance and dialogical status can also be applied to the negotiation part of a dialogue. These mechanisms will then be used to ensure that as long as a persuasion move is legal, no negotiation move can be made: thus the protocol will capture the idea of embedding persuasion in negotiation.

3.1 The combination

First the combined communication language L_c is defined in Table 3. As can be seen, the negotiation language is reformulated in the format of Section 2.3 by dividing the "Replies" of Table 1 into surrendering replies ($accept(\varphi)$) and attacking replies (all other replies). Next a new attacking reply is added, viz. *why-reject*(φ) as a reply to *reject*(φ). The only possible reply to this new locution other than a withdrawal is with the persuasion locution $claim(\neg\varphi)$. The use of this reply induces a shift from a negotiation to a persuasion subdialogue.

Table 3. Speech acts and replies in L_c.

Acts	Attacks	Surrenders
negotiation		
$request(\varphi)$	$offer(\varphi')$ *withdraw*	
$offer(\varphi)$	$offer(\varphi'))$ $(\varphi \neq \varphi')$ $reject(\varphi)$ *withdraw*	$accept(\varphi)$
$reject(\varphi)$	$offer(\varphi')$ $(\varphi \neq \varphi')$ $why\text{-}reject(\varphi)$ *withdraw*	
$accept(\varphi)$		
$why\text{-}reject(\varphi)$	$claim(\neg\varphi)$ *withdraw*	
withdraw		
persuasion		
$claim(\varphi)$	$why(\varphi)$	$concede(\varphi)$
$why(\varphi)$	$argue(A)$ $(conc(A) = \varphi)$	$retract(\varphi)$
$argue(A)$	$why(\varphi)$ $(\varphi \in prem(A))$ $argue(B)$ (B defeats A)	$concede(\varphi)$ $(\varphi \in prem(A)$ or $\varphi = conc(A))$
$concede(\varphi)$		
$retract(\varphi)$		

Next, in order to specify the combined protocol, the notion of negotiation moves must be adapted to fit the format of Definition 1 (which we leave implicit). The combined protocol is then defined as follows.

Definition 4. *(The protocol P for L_c.) For all dialogues d and moves m it holds that $m \in P(d)$ if and only if m satisfies all of the following rules.*

- R_1*: m satisfies $R_1 - R_8$ of Definition 3 but where in R_2, L_c^p is replaced by L_c and in R_5, claim(φ) is replaced by request(φ);*
- R_2*: If $s(m) =$ offer(φ) and $s(m_1) =$ request(φ') then $\{\varphi, \varphi'\}$ is consistent and φ contains at least the same issues as φ';*
- R_3*: If $s(m) =$ offer(φ) then of no $m' \in d$, $s(m') =$ offer(φ);*
- R_4*: If $s(m) =$ accept(φ) then φ contains no variables;*
- R_5*: If m is a negotiation locution then m replies to the most recent target to which a reply is legal;*
- R_6*: If m is a negotiation locution then there is no move $m' \in P(d)$ such that $s(m')$ is a persuasion locution;*
- R_7*: If $s(m) =$ offer(φ) then $C_s(d) \cup \{\varphi\}$ and $C_{\overline{s}}(d) \cup \{\varphi\}$ are consistent.*

Rule R_1 generalises the general structure of the persuasion protocol to the combined protocol and says that each combined dialogue starts with a request for an offer. Rules $R_2 - R_4$ formalise the negotiation protocol rules of [4] that are not implied by R_1 (see also below). Rule R_5 prevents unnecessary negotiation backtracking moves. Finally, rules R_6 and R_7 perform a key role in the embedding of persuasion in negotiation. R_6 enforces that the relation between the negotiation and persuasion parts of dialogues is one of embedding of the latter in the former (cf. [17]): as long as a persuasion move is legal, no negotiation move is legal. And R_7 formalises the intuition that offers need to respect the reasons for rejection given by the other party when these reasons have been successfully defended in an embedded persuasion dialogue.

Rule R_7 is justified by the following property of the persuasion protocol of [5]: under some plausible assumptions on the contents of arguments a *retract*(t) move in reply to a challenge of the initial claim is always legal. Then by R_6 of the persuasion protocol, which requires retractions to be successful, a player who has defended a rejection with a *claim*(t) move in a terminated persuasion dialogue is committed to t only if he has won the persuasion dialogue about t.

The turntaking rule of the combined system is the same as for persuasion. Given L_c, this rule implies that just as in Section 2.1 the turn shifts after each negotiation move except after an *accept* move, which terminates a dialogue.

Finally, the new commitment rules need to be defined. In fact, they are the same as for persuasion moves in Section 2.3. The effects that negotiation moves have on the players' commitments are irrelevant as long as a dialogue has not terminated, since an offer commits the offeree to an action only after the offer has been accepted: so checking compliance with negotiation commitments lies outside the negotiation dialogue in which the commitment was incurred.

Note that the new system completely preserves the original persuasion system and as much as possible preserves the original negotiation system. Above we already noted that turntaking in the negotiation part is still the same. Furthermore, backtracking from negotiation moves (which was impossible in the original system) is legal in two cases only: if the one who challenges a rejection loses the resulting persuasion dialogue, s/he must move an alternative reply to the rejection, and if the other party loses such a persuasion dialogue, s/he must move a counteroffer or withdrawal in reply to the rejected offer.

3.2 Properties of the combined protocol

The main property of the new protocol is about the maximum number of negotiation moves needed to reach a certain agreement.

Proposition 2. *For any proposal φ the maximum length of a negotiation dialogue to end with acceptance of φ is never higher and sometimes lower in the system of Section 3 than in the system of Section 2.1.*

proof: This follows from the fact that the only effect of a terminated persuasion dialogue on an embedding negotiation dialogue is that it may make offers illegal since they do not respect the commitments of the other agent. Thus the number of legal offers in a negotiation according to Section 3 is never higher and sometimes lower than in a negotiation according to Section 2.1.

Since our persuasion protocol is not guaranteed to terminate, the same holds for our combined protocol. However, on the assumption that a persuasion dialogue always terminates, Proposition 2 implies that the 'success' result on the negotiation protocol proven by [4] still holds for our combined protocol: if the set of possible outcomes is finite then any negotiation is guaranteed to terminate with a withdraw or an accept.

4 An example

We next illustrate our new protocol by extending our example from Section 2.1 with an embedded persuasion dialogue. For simplicity we paraphrase the contents of the arguments and we do not formally distinguish beliefs, desires and intentions, as is done in e.g. [18,19]. To illustrate the use of defeasible inference rules, some arguments are assumed to be constructed with presumptive argumentation schemes from [20]. In [21] it is discussed how such schemes can be formalised as defeasible inference rules and their critical questions as pointers to undercutters. Elementary inferences within arguments are paraphrased as *conclusion since premises*. All moves in the dialogue except proponent's last four moves reply to their immediate predecessor.

P_1: $request(brand = ? \wedge price = ?)$
O_2: $offer(brand = peugeot \wedge price = 10000)$
P_3: $reject\ (brand = peugeot \wedge price = 10000)$
Olga now exploits the additional features of the protocol by asking Paul why he rejected the offer.
O_4: $why\text{-}reject(brand = peugeot \wedge price = 10000)$
Paul now meets Olga's challenge of his rejection so that the negotiation shifts into a persuasion. All persuasion moves below until P_{14} reply to their immediate

predecessor.
P_5: $claim(\neg\ (brand = peugeot \wedge price = 10000))$
Paul is now committed to the content of his claim.
O_6: $why\ (\neg\ (brand = peugeot \wedge price = 10000))$
P_7: $argue\ (\neg\ (brand = peugeot \wedge price = 10000)$ *since* $\neg\ brand = peugeot$; *this in turn is so since* $(brand = peugeot \rightarrow brand = french)$ *and* $\neg\ brand = french)$
It turns out that Paul rejected the offer since a Peugeot is a French car and he does not want French cars. If Olga now simply concedes Paul's claim as an alternative reply to P_7, the persuasion dialogue terminates and the negotiation is resumed. Then Olga cannot reply to P_3 in the same way as in section 2.1 by offering another french car. Olga could offer a non-French car (as in O_6 in section 2.1) but she chooses to try to persuade Paul that he is wrong in not wanting a French car and she therefore challenges Paul's second premise.
O_8: $why\ (\neg\ brand = french)$
P_9: *argue* ($\neg$ *brand* = *french since having french cars is bad; this is so since french cares are unsafe and having an unsafe care usually is bad.*)
Paul defends his second premise with an argument from (bad) consequences.
O_{10}: *why* (*french cars are unsafe*)
P_{11}: *argue* (*french cars are unsafe since car magazine mycar says so and mycar are experts about cars*)
This is a defeasible argument based on the argumentation scheme from expert opinion: "what experts say is normally true".
O_{12}: *argue* (*magazine mycar is biased since magazine mycar is german and german car magazines are usually biased against french cars*)
Using a default rule, Olga constructs an undercutter of the argument from expert opinion, namely that this expert is biased.
P_{13}: *concede* (*magazine mycar is german*)
P_{14}: *concede* (*german car magazines are usually biased against french cars*)
Even though Paul has conceded the premises of Olga's undercutter, he can still move a counterargument, since the argument is defeasible because it uses a default rule. Paul chooses to rebut the undercutter, using another default rule.
P_{15}: *argue* ($\neg$ *magazine mycar is biased since magazine mycar has a very high reputation and car magazines with high reputation usually are not biased*)
Note that P_{14} is a second and P_{15} a third reply to O_{12}.
O_{16}: *why* (*magazine mycar has a very high reputation*)
Let us assume that Paul now realises that he has no plausible way to defend his premise that the car magazine has a high reputation. At this point, all of Olga's persuasion moves are relevant targets for Paul. He could, for instance, move another rebuttal of Olga's undercutter, or another argument why French cars are not safe or why he does not want french cars. But suppose that Paul sees no plausible way of doing so and instead retracts that he does not want French cars by moving an alternative reply to O_8 and then retracting his main claim as a second reply to O_6.
P_{17}: $retract\ (\neg\ brand = french)$
P_{18}: $retract\ (\neg\ (brand = peugeot \wedge price = 10000))$

Now Paul has no legal persuasion moves any more since all targets have become irrelevant: since Paul has surrendered to O_6, his main claim P_5 cannot be changed from *out* to *in*. So the persuasion dialogue terminates and the negotiation resumes with Olga to move after P_2. Since with P_{18} Paul has ended his commitment to his main claim, Olga is now allowed to offer another French car, perhaps even a Peugeot for a lower price. The negotiation could now continue as in Section 2.1 with move O_4.

It is instructive to construct the dialectical graph of arguments and counterarguments exchanged by Paul and Olga during the persuasion dialogue ($p \rightsquigarrow q$ reads as "if p then usually q").

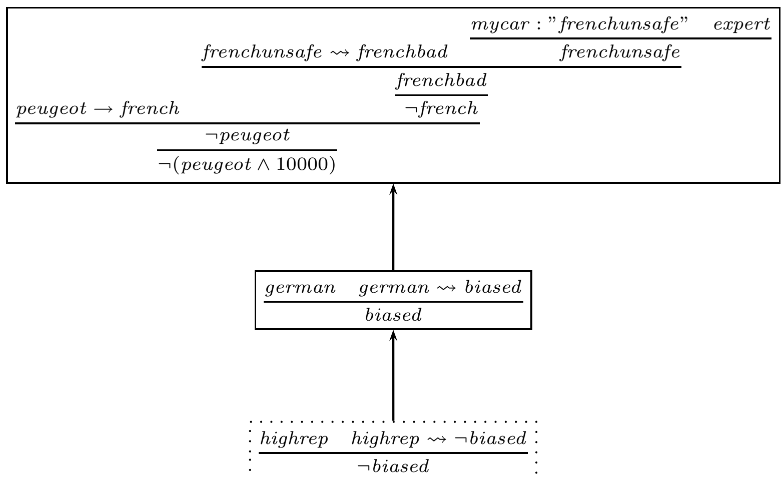

Figure 2. The dialectical graph

The graph contains a simple argument game according to the proof theory of the underlying logic. Since on the basis of the information exchanged during the persuasion dialogue no other counterarguments to one of these three arguments can be constructed, the graph is actually a proof that, on the basis of this information, the proposition $\neg(peugeot \wedge 10000)$ is justified. However, the last argument in the graph has one challenged premise, viz. *highrep*, so this argument is not defended (indicated by the dotted box). The defended part of the graph is instead a proof that $\neg(peugeot \wedge 10000)$ is not justified on the basis of all *defended* information.

5 Conclusion

In this paper we have presented a protocol for negotiation with embedded argumentation that has a social semantics. In doing so, we have exploited the general format of [6,5] of dialogue systems. In the resulting dialogue game reasons for rejections can be asked and, when given, can constrain further offers unless the offering agent can persuade the rejecting agent that the reason is not tenable. Thus a negotiation is sometimes more efficient since offers that will certainly be rejected can be avoided, and it is sometimes of higher quality since flawed reasons can be revised. The persuasion protocol is flexible in that it allows for different underlying logics, for alternative replies and for postponing replies, sometimes even indefinitely. Yet a strong focus of dialogues is maintained through the requirement of relevance.

We know of one earlier protocol that allows for persuasion dialogues about rejections, viz. [3]. It was a source of inspiration for the present work but there are reasons for further development. The first is that the protocol does not have a social semantics, since whether an agent is allowed to assert a claim or an argument partly depends on their internal mental state. Also, arguments have to be classical propositional proofs from a consistent set of premises so that, for instance, the use of presumptive argument schemes or undercutting counterarguments is not supported. A further limitation is that the dialectical aspects of the underlying logic are only used internally by an agent, to verify whether they have an acceptable argument for a claim in their (possibly inconsistent) knowledge base. By contrast, in our protocol the dialectical role of each argument in a dialogue is made explicit, as illustrated by Figure 2. Finally, the protocol of [3] only weakly maintains focus of dialogues, allowing, for example, dialogues like P_1: *claim*(p), O_2: *why*(q).

In future research the present protocol should be combined with relevant other work. For instance, [18,19] define a rich topic language in which the beliefs, desires and intentions of agents can be distinguished and reasoned about, allowing negotiating agents to produce and attack several interesting types of arguments. Since we have partially abstracted from the nature of the persuasion topic language and defeasible inference rules, our dialogue system can be instantiated with this work. We also aim to study the interaction of the present protocol with agent designs and strategies, as, for instance, done in [8] for several dialogue types. Finally, we aim to include other forms of argument-based negotiation, such as arguments why a proposal should be accepted ([18,19]).

Acknowledgements

Henry Prakken was partially supported by the EU under IST-FP6-002307 (ASPIC). The authors thank Birna van Riemsdijk for her useful comments on a draft of this paper.

References

1. Rahwan, I., Ramchurn, S., Jennings, N., McBurney, P., Parsons, S., Sonenberg, L.: Argumentation-based negotiation. The Knowledge Engineering Review **18** (2003) 343–375

2. Singh, M.: Agent communication languages: rethinking the principles. IEEE Computer **31** (1998) 40–47
3. Amgoud, L., Parsons, S., Maudet, N.: Arguments, dialogue, and negotiation. In: Proceedings of the Fourteenth European Conference on Artificial Intelligence. (2000) 338–342
4. Wooldridge, M., Parsons, S.: Languages for negotiation. In: Proceedings of the Fourteenth European Conference on Artificial Intelligence. (2000) 393–400
5. Prakken, H.: Coherence and flexibility in dialogue games for argumentation. Technical Report UU-CS-2005-021, Department of Information and Computing Sciences, Utrecht University, Utrecht (2005)
6. Prakken, H.: On dialogue systems with speech acts, arguments, and counterarguments. In: Proceedings of the 7th European Workshop on Logic for Artificial Intelligence (JELIA'2000). Number 1919 in Springer Lecture Notes in AI, Berlin, Springer Verlag (2000) 224–238
7. Walton, D., Krabbe, E.: Commitment in Dialogue. Basic Concepts of Interpersonal Reasoning. State University of New York Press, Albany, NY (1995)
8. Parsons, S., Wooldridge, M., Amgoud, L.: Properties and complexity of some formal inter-agent dialogues. Journal of Logic and Computation **13** (2003) 347-376.
9. Prakken, H., Vreeswijk, G.: Logics for defeasible argumentation. In Gabbay, D., Günthner, F., eds.: Handbook of Philosophical Logic. Volume 4. Second edn. Kluwer Academic Publishers, Dordrecht/Boston/London (2002) 219–318
10. Dung, P.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n–person games. Artificial Intelligence **77** (1995) 321–357
11. Prakken, H., Sartor, G.: Argument-based extended logic programming with defeasible priorities. Journal of Applied Non-classical Logics **7** (1997) 25–75
12. Amgoud, L., Cayrol, C.: A model of reasoning based on the production of acceptable argument. Annals of Mathematics and Artificial Intelligence **34** (2002) 197–216
13. Pollock, J.: Cognitive Carpentry. A Blueprint for How to Build a Person. MIT Press, Cambridge, MA (1995)
14. Dung, P.: Logic programming as dialog game. Unpublished paper, Division of Computer Science, Asian Institute of Technology, Bangkok (1994)
15. Barwise, J., Moss, L.: Vicious Circles. Number 60 in CSLI Lecture Notes. CSLI Publications, Stanford, CA (1996)
16. Mackenzie, J.: Question-begging in non-cumulative systems. Journal of Philosophical Logic **8** (1979) 117–133
17. McBurney, P., Parsons, S.: Games that agents play: A formal framework for dialogues between autonomous agents. Journal of Logic, Language and Information **13** (2002) 315–343
18. Kraus, S., Sycara, K., Evenchik, A.: Reaching agreements through argumentation: a logical model and implementation. Artificial Intelligence **104** (1998) 1–69
19. Parsons, S., Sierra, C., Jennings, N.: Agents that reason and negotiate by arguing. Journal of Logic and Computation **8** (1998) 261–292
20. Walton, D.: Argumentation Schemes for Presumptive Reasoning. Lawrence Erlbaum Associates, Mahwah, NJ (1996)
21. Bex, F., Prakken, H., Reed, C., Walton, D.: Towards a formal account of reasoning about evidence: argumentation schemes and generalisations. Artificial Intelligence and Law **12** (2003) 125–165