Argument & Computation
Vol. 00, No. 00, Month 2009, 1–18

Formal Reconciliatory Dialogue Based on Shift from Forward to Backward Deliberation†

Hiroyuki Kido[a*] and Federico Cerutti[b]

[a] *The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan*
[b] *Cardiff University, Queen's Buildings, 5 The Parade, Cardiff CF24 3AA, UK*

(*v1.0 released Feb 2015*)

Desire conflicts arise in several real-world contexts. In this paper, we propose a mixed deliberation dialogue for reconciliation. A mixed deliberation dialogue is defined as a combination of *forward* and *backward* deliberation dialogues with respective goals are *subordinate* and *superordinate* desires of a given desire. This research and the introduction of mixed deliberation dialogue have been motivated by Kowalski and Toni's reconciliatory scenario. We show that an instantiation of a mixed deliberation dialogue implements key parts of Kowalski and Toni's reconciliatory solution. We also prove the correctness of mixed deliberation dialogues.

Keywords: Formal dialogue, deliberation, dialogue shift, reconciliation, defeasible inference rules

1. Introduction

Dialogue theory encompasses various types of descriptive and formal studies, aimed at various purposes, on the structure of dialogues (van Eemeren et al. 1996). Hintikka's game-theoretic semantics (Hintikka 1968) and Lorenzen's dialogue logic (Lorenzen 1961) explore semantics of language. Hamblin's formal dialectics (Hamblin 1970) explores descriptive or formal dialogue systems. Because formal dialogue systems can give agents rational interaction and computation mechanisms under uncertain, incomplete, inconsistent, subjective, and distributed information, they have received attention from researchers working on formal argumentation (Fan and Toni 2012, Kok et al. 2010, Prakken 2005, 2006, Wells and Reed 2006).

However, little work has been done for dialogue systems for reconciling conflict not only by searching for means of satisfying either all or part of given desires, but also by searching for means for satisfying their underlying desires behind the given ones. Kowalski and Toni (Kowalski and Toni 1994) first presented arguments for the necessity of reconciliation in the context of argumentation.

What is interesting about their scenario is that neither the generalized goal nor reconciliatory solution is obtainable merely by just choosing one of the given alternatives based on utilities or preferences, i.e., quantitative measures, but their scenario requires some sort of qualitative measure to shift to an underlying desire. Based on the scenario, they outlined what a generalized goal and a reconciliatory solution are. However, an open question is how one should have a dialogue to reach the generalized goal and the reconciliatory solution. These observations motivate

†This paper is a revised full version of the conference paper (Kido and Cerutti 2014).
*Corresponding author. Email: kido@sys.t.u-tokyo.ac.jp

ISSN: 1946-2166 print/ISSN 1946-2174 online

DOI: 10.1080/1946216YYxxxxxxxx
http://www.informaworld.com

us to formalize reconciliatory dialogues as consisting of *forward* and *backward* deliberation dialogues.

This paper contributes to the state-of-the art of studies of formal dialogue and argumentation by handling the processes leading from conflict detection to justification of reconciliation in terms of a series of dialogues. Particularly, this paper gives underlying dialogue and inference principles behind reconciliation. Furthermore, we address Kowalski and Toni's academically acknowledged scenario that cannot be solved simply by taking advantage of utilities or preferences assumed in many formal dialogue-based and argumentation-based approaches.

This paper is organized as follows. Section 2 motivates the research referring to Kowalski and Toni's scenario: a running example is used throughout the paper to describe our proposal. Section 3 presents some preliminary notions that are used widely to define dialogue goals in Section 4 and dialogue protocols in Section 5. In Section 6 we proved the correctness of the dialogues with respect to their goals. Section 7 discusses related work and Section 8 concludes the paper. Proofs of results are in Appendix A.

2. Motivation

We consider the following realistic reconciliatory scenario demonstrating the importance of goal generalization.

Example 1. (Kowalski and Toni 1994) In a recent head-of-sections committee meeting in our Department, we discussed the composition of a new resources committee. Two conflicting arguments were put forward. The Director of Administration argued that, in the interests of efficiency, the members of the new committee should consist of himself and the other principal administrative officers of the Department. The Director of Research argued, in opposition to him, that, in the interests of democracy, the committee should also contain members elected by the Department.

During the course of the discussion it became clear that the two sides were focusing on different assumptions about the purpose of the new committee: the Director of Administration on its purely administrative function, and the Director of Research on its presumed policy making nature. These two assumptions could be viewed as conflicting solutions to the more general goals of deciding, on the one hand, which group should administer resources, and on the other hand, which group should make policy about resources.

By focusing on the more general goals, it was possible to identify a new solution which was acceptable to both parties: the resources committee will administer resources, whereas the head-of-sections committee will make policy about resources. In the interests of efficiency, the members of the resources committee will consist of administrative officers only. In the interests of democracy, the head-of-sections committee will represent the views and interests of the various Department sections on matters concerning policy about the allocation of resources.

In this scenario, neither a generalized goal nor a reconciliatory solution is obtainable merely by just choosing a given alternative based on utility or preference. An open question is how one should have a dialogue to reach the generalized goal and the reconciliatory solution.

Therefore, we formalize reconciliatory dialogues as consisting of *forward* and *backward* deliberation dialogues. We show how shifting between *forward* and *backward* deliberation dialogue highlights the linkage with more general goals. On one hand, the Director of Administration desires *efficiency*. On the other hand, the

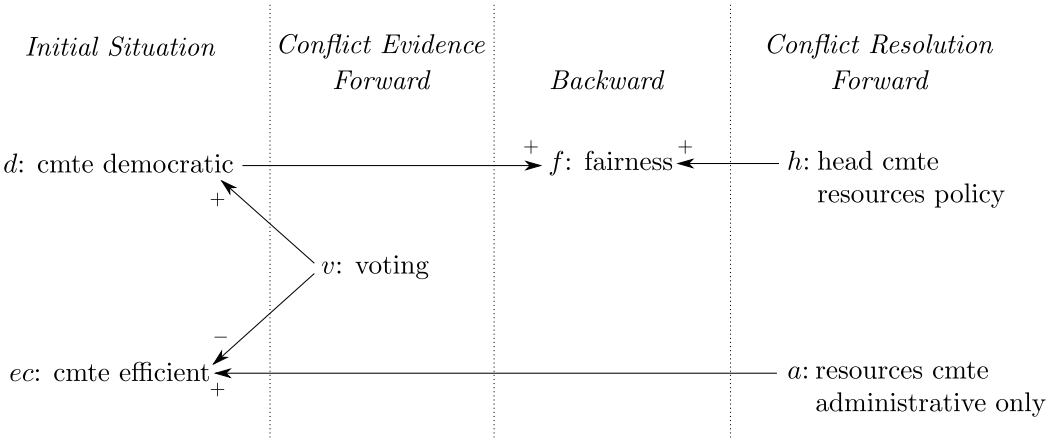

Figure 1. Rough causal relations behind Kowalski and Toni's reconciliatory scenario

Director of Research wants to promote *democracy* via *voting* although this demotes *efficiency* — *forward deliberation*. Then, because one outcome of adopting *democracy* is *fairness*, and because there is no reason in favor of not desiring *fairness*, we can consider it a "desirable" outcome of the *democracy* — *backward deliberation*. Finally, shifting back to a *forward* deliberation dialogue, the two more general goals become evident, because *efficiency* will be improved if the resources committee comprises administrative officers only, and *fairness* will be ensured if the head committee will take care of making policies related to resources.

In Figure 1, we show rough causal relations behind their scenario with some additional information where nodes respectively represent statements and arrows, from node x to node y, attached with $+$ (resp. $-$) represent x promotes (resp. demotes) y. For the discussion, we expand Kowalski and Toni's scenario.

Example 1 (continued). According to Figure 1:

- d represents "the resources committee is democratic";
- ec represents "the resources committee is efficient";
- v represents "the resources committee is elected by vote";
- f represents "the policy is fair";
- h represents "the head committee makes policy about resources";
- a represents "the resources committee is composed by administrative officers only".

3. Preliminaries

We use Dung's theory of acceptability semantics (Dung 1995) that reformulates consequence notions of nonmonotonic logics. The semantics is defined on a pair $AF = \langle AR, attacks \rangle$, called an abstract argumentation framework, where AR is a set of arguments and $attacks$ is a binary relation on AR, i.e., $attacks \subseteq AR \times AR$.

Definition 1. (Dung 1995) Let $AF = \langle AR, attacks \rangle$ be an abstract argumentation framework, $S \subseteq AR$ and $a \in AR$.

- S is *conflict-free* iff no $a, b \in S$ exists such that a attacks b, i.e., $(a, b) \notin attacks$.
- a is *acceptable* with respect to S iff, for all $b \in AR$, if b attacks a then there is $c \in S$ such that c attacks b .
- The *characteristic function*, $F_{AF} : Pow(AR) \rightarrow Pow(AR)$, is defined as $F_{AF}(S) = \{a \in AR \mid a \text{ is acceptable with respect to } S\}$.

- S is the *grounded extension* iff it is the least fixed point of F_{AF}.

Given an argumentation framework $AF = \langle AR, attacks \rangle$, we say that an argument $a \in AR$ is *justified* in AF iff a is in the grounded extension of AF, and *overruled* otherwise.

In the following, we consider argumentation frameworks instantiated from $\mathcal{L}_0$, a language of modal propositional logic with single modal operator D representing the operator "it is desirable that". $\mathcal{L}_0$ is closed under truth-functional operations. Consequently, if $a \in \mathcal{L}_0$, then $\neg a \in \mathcal{L}_0$ and if $a, b \in \mathcal{L}_0$, then $a \vee b, a \wedge b, a \rightarrow b \in \mathcal{L}_0$ etc. $\mathcal{L}_0$ conforms to the axiomatic system KD in which $D(a \rightarrow b) \rightarrow Da \rightarrow Db$ and $Da \rightarrow \neg D \neg a$ are axioms. Also, $\mathcal{L}_1$ consists of so-called defeasible conditionals, or defaults. They commonly have the forms "$a \Rightarrow b$" where a is a conjunction of literals, i.e., atomic propositions or their negation, in $\mathcal{L}_0$ and b is a literal in $\mathcal{L}_0$, and mean that if a is the case, then b is normally the case. Operator D is assumed not to appear in defeasible conditionals. We assume a fixed, but arbitrary theory $T \subseteq \mathcal{L}_0 \cup \mathcal{L}_1$.

Example 1 (continued). To illustrate our proposal, we consider the following additional propositions:

- e represents "the policy is effective";
- m represents "the resources committee makes policy about resources";
- r represents "the head of sections committee represents various views of interests"

Moreover, let us consider the following rules, which provide the casual relations depicted in Figure 1:

$$v \Rightarrow d; \qquad v \Rightarrow \neg ec; \qquad a \Rightarrow ec; \qquad d \wedge m \Rightarrow f;$$
$$d \wedge m \Rightarrow \neg e; \qquad ec \wedge m \Rightarrow e; \qquad h \wedge r \Rightarrow f.$$

Any rule of inference that is not valid with respect to modal logic KD is called a defeasible inference rule, represented by $\rightsquigarrow$. More precisely, a rule of inference is not valid if it can derive a formula that is not a theorem in KD. Intuitively, an inference is not valid if it is not deductive in terms of KD. We use letters $a, b, c, ...$ of the alphabet to represent literals in $\mathcal{L}_0 \cup \mathcal{L}_1$, lower-case Greek letters $\alpha, \beta, \gamma, ...$ to represent their metavariables of them and $A, B, C, ...$ to represent their sequences.

Definition 2. (Bench-Capon and Prakken 2006) Positive forward practical syllogism, denoted by $PFPS$, and negative forward practical syllogism, denoted by $NFPS$, are defined as follows, respectively[1].

$$PFPS : D\alpha, \gamma, \beta \wedge \gamma \Rightarrow \alpha \rightsquigarrow D\beta$$
$$NFPS : D\alpha, \gamma, \beta \wedge \gamma \Rightarrow \neg\alpha \rightsquigarrow D\neg\beta$$

The positive one states intuitively that if one believes that α is desirable (e.g. "democratic cmte" d), γ is the case and if β (e.g. "voting" v) is realized under γ is the case then α is satisfied, then one defeasibly concludes that β is desirable ($Dd, v \Rightarrow d \rightsquigarrow Dv$). The negative one, however, concludes that β is undesirable from a different premise stating that if β is realized under γ is the case, then α

[1]The authors originally call them positive and negative practical syllogisms, respectively, and the conclusion part of the negative one is $\neg D\beta$.

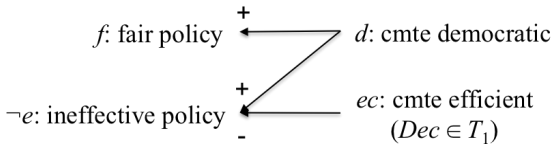

Figure 2. Rough causal relations behind the superordinate desire in Example 1.

is frustrated. We describe them as FPS without distinction. We often silently use abbreviated forms $D\alpha, \beta \Rightarrow \alpha \rightsquigarrow D\beta$ and $D\alpha, \beta \Rightarrow \neg\alpha \rightsquigarrow D\neg\beta$.

Definition 3. Let $\Sigma \subseteq T$, $p_i \in \mathcal{L}_0 \cup \mathcal{L}_1$. A sequence $A = p_1, p_2, ..., p_n$ is an *argument* from Σ to p_n iff (1), for all $i (1 \leq i \leq n)$, $p_i \in \Sigma$ (base case) or p_i is derived from preceding formulae $p_j (j < i)$ by application a rule of inference, and (2), for all $i (1 \leq i \leq n)$, the sequence obtained by eliminating p_i from A does not satisfy (1).

The first condition states that an argument is a derivation from Σ using strict and defeasible inference rules. The second assures that an argument consists of a minimal number of formulae. As described herein $\Sigma \vdash_x p_n$ denotes an argument from Σ to p_n where rules of inference are restricted to only x. For example, $\Sigma \vdash_{FPS} Da$ represents that Da is derived from Σ by application only positive or negative forward practical syllogisms zero or more times.

4. Formal Goals of Deliberation Dialogues

4.1. *Superordinate and Subordinate Desires*

This section provides formal definitions of dialogue goals. Given a desirable outcome or desire, we use the term *superordinate desire* to refer to a desire such that, once it is assumed, it can be a rationale for desiring a given desire, but it cannot be a rationale for not desiring a given desire. Based on the recognition that the practical syllogisms give fundamental inference mechanisms for practical reasoning, we give a formal definition of superordinate desires as follows.

Definition 4. Let $Dg, Dh \in \mathcal{L}_0$. Dh is a *superordinate desire* of Dg in T iff $\Sigma_1 \subseteq T$ exists such that $\Sigma_1 \cup \{Dh\} \vdash_{FPS} Dg$ and no $\Sigma_2 \subseteq T$ exists such that $\Sigma_2 \cup \{Dh\} \vdash_{FPS} D\neg t$, for all $Dt \in T \cup \{Dg\}$.

Definition 4 states that a superordinate desire Dh derives Dg by application of only forward practical syllogisms $\vdash_{FPS}$, but no negation of any desire in $T \cup \{Dg\}$ can be derived by their application. It states that h would be desirable because, once Dh is assumed, it can be a rationale for Dg. Definition 4 is weaker than the definition replacing Σ to T because it permits $T \vdash_{FPS} Dg$, i.e., the situation in which Dg is derived from T without using Dh. Moreover, in general, a superordinate desire is not intrinsically a desire existing in theory T, but it is derived by defeasible inferences.

Example 1 (continued). Df is a superordinate desire of Dd in $T = \{d \wedge m \Rightarrow f, d \wedge m \Rightarrow \neg e, ec \wedge m \Rightarrow e, Dec, m\}$. However, $D\neg e$ is not because of the following reason. Once $D\neg e$ is assumed, Dd is derived from $\Sigma_1 = \{d \wedge m \Rightarrow \neg e, m\} (\subseteq T)$, i.e., $\Sigma_1 \cup \{D\neg e\} \vdash_{FPS} Dd$. However, once $D\neg e$ is assumed, $D\neg ec$ is derived from $\Sigma_2 = \{ec \wedge m \Rightarrow e, m\} (\subseteq T)$, i.e., $\Sigma_2 \cup \{D\neg e\} \vdash_{FPS} D\neg ec$. Figure 2 shows rough causal relations behind this superordinate desire. One can see that Dd is derived from the assumption Df using FPS. Although it can also be derived from the assumption $D\neg e$, the assumption results in a derivation of $D\neg ec$ that conflicts

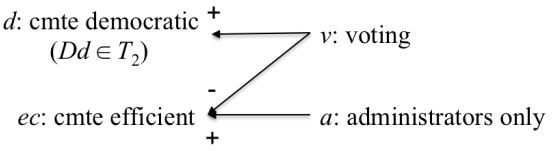

Figure 3. Rough causal relations behind the subordinate desire in Example 1.

with $Dec \in T$.

We introduce a notion of *subordinate desires*. Given a desire, we use the term *subordinate desire* to refer to a desire such that if once it is realized then it satisfies at least one of the given desires, but it frustrates no desire in them. Namely, a subordinate desire of a desire is desirable as a means of satisfying the desire, but not as a means of frustrating any other desire.

Definition 5. Let $Dg, Dh \in \mathcal{L}_0$. Dg is a *subordinate desire* of Dh in T iff there is $\Sigma_1 \subseteq T$ such that $\Sigma_1 \cup \{Dh\} \vdash_{FPS} Dg$ and there is no $\Sigma_2 \subseteq T$ such that $\Sigma_2 \cup \{Dt\} \vdash_{FPS} D\neg g$, for all $Dt \in T \cup \{Dh\}$.

Definition 5 states that the subordinate desire Dg is derived by application only of forward practical syllogisms, but no desire in $T \cup \{Dh\}$ derives its negation $D\neg g$ by application of them. It states that g is desirable as a means for satisfying h without frustrating any desire in $T \cup \{Dh\}$. As with superordinate desires, Definition 5 is weaker than the definition replacing Σ to T. In general, a subordinate desire is not intrinsically a desire existing in theory T.

Example 1 (continued). Da is a subordinate desire of Dec in $T = \{v \Rightarrow d, v \Rightarrow \neg ec, a \Rightarrow ec, Dd\}$. However, $D\neg v$ is not because, given $\Sigma_1 = \{v \Rightarrow \neg ec\}$, although $\Sigma_1 \cup \{Dec\} \vdash_{FPS} D\neg v$, but given $\Sigma_2 = \{v \Rightarrow d\}$, $\Sigma_2 \cup \{Dd\} \vdash_{FPS} Dv$. Figure 3 shows rough causal relations behind this subordinate desire. It might be readily apparent that Da is derived from the assumption Dec using FPS. Although $D\neg v$ can also be derived from Dec, it conflicts with Dv derived from $Dd \in T$ using FPS.

4.2. *Reconciliatory Desires as a Combination of Superordinate and Subordinate Desires*

Finally, we introduce the notion of *reconciliatory desires* defined by combining superordinate desires and subordinate desires.

Definition 6. Let $Dg, Dh \in \mathcal{L}_0$. Dg is a *reconciliatory desire* of Dh in T iff $Di \in \mathcal{L}_0$ exists such that Di is a superordinate desire of Dh in T and Dg is a subordinate desire of Di in T.

Definition 6 states that g is desirable because it does not frustrate any desire in $T \cup \{Di\}$, but satisfies i. Here, i is regarded as desirable because, once it is assumed, it gives a rationale for desiring h. Note that subordinate and superordinate desires are both special cases of reconciliatory desires.

Example 1 (continued). $T = \{d \wedge m \Rightarrow f, d \wedge m \Rightarrow \neg e, ec \wedge m \Rightarrow e, m, Dec, h \wedge r \Rightarrow f, r\}$. Dh is a reconciliatory desire of Dd in T because Df is a superordinate desire of Dd, and Dh is a subordinate desire of Df.

4.3. *Backward Practical Syllogisms*

We introduce a *backward* version of the practical syllogism. This inference pattern, differently from the traditional forward practical syllogism, represents the result of a critical thinking approach. Indeed, presuming that α is a desirable outcome ($D\alpha$), and assuming it is true that $\alpha \Rightarrow \beta$. Therefore, if α is accepted as desirable, then β, because it is a material implication deriving from α, should be considered. What the backward practical syllogism is doing here is to suggest exploration of the world by defeasibly assuming that $D\beta$ holds too. Clearly, this is the case only if there is no evidence of the contrary — i.e. that $D\neg\beta$ holds. As we will see in Section 5, such contrary is formally handled by interaction of inferences in a dialogue, instead of assuming *negation as failure* $\sim D\neg\beta$ meaning that each attempt to prove $D\neg\beta$ fails.

Definition 7. Positive backward practical syllogism, denoted by $PBPS$, and negative backward practical syllogism, denoted by $NBPS$, are defined as follows.

$$PBPS : D\alpha, \gamma, \alpha \wedge \gamma \Rightarrow \beta \rightsquigarrow D\beta$$
$$NBPS : D\alpha, \gamma, \neg\alpha \wedge \gamma \Rightarrow \beta \rightsquigarrow D\neg\beta$$

The positive backward practical syllogism intuitively states that if one believes that α is desirable (e.g. "cmte democratic" d), γ is the case (e.g. "the resources committee makes policy about resources" m) and if α is realized under γ is the case then β is realized (e.g. "fairness" f). Then one defeasibly concludes that β might be desirable ($Dd, m, d \wedge m \Rightarrow f \rightsquigarrow Df$). The negative one, however, concludes that β is undesirable from a different premise stating that if α is not realized under γ is the case then β is realized. We describe them as BPS without distinction.

An application of BPS sometimes derives a false conclusion. The following example shows that how such false conclusions can be withdrawn by an interaction of BPS.

Example 1 (continued). The following is an application of a backward practical syllogism.

$$Dd, m, d \wedge m \Rightarrow \neg e \rightsquigarrow D\neg e$$

Namely, one believes that a democratic resources committee is desirable (i.e., Dd), the resources committee makes policy about resources (i.e., m) and if d is realized under m is the case then it demotes efficiency of the policy (i.e., $d \wedge m \Rightarrow \neg e$). From these beliefs, the backward practical syllogism defeasibly derives the belief that an ineffective policy is desirable (i.e., $D\neg e$).

This is intuitively a false conclusion. As we will see in Section 5, the conclusion is withdrawn by agent's belief De or the following another application of a backward practical syllogism.

$$Dec, m, ec \wedge m \Rightarrow e \rightsquigarrow De$$

Namely, one believes that an efficient resources committee is desirable (i.e., Dec), the resources committee makes policy about resources (i.e., m) and if ec is realized under m is the case then it promotes efficiency of the policy (i.e., $ec \wedge m \rightsquigarrow e$). From these beliefs, the backward practical syllogism defeasibly derives the belief that an effective policy is desirable.

Finally, in this section, we show the fact that backward practical syllogisms

redefine superordinate desires originally defined by forward practical syllogisms.

Proposition 1. Let $Dg, Dh \in \mathcal{L}_0$. Dh is a superordinate desire of Dg in T iff there is $\Sigma_1 \subseteq T$ such that $\Sigma_1 \cup \{Dg\} \vdash_{BPS} Dh$ and there is no $\Sigma_2 \subseteq T$ such that $\Sigma_2 \cup \{Dt\} \vdash_{BPS} D\neg h$, for all $Dt \in T \cup \{Dg\}$.

5. Formal Protocols of Deliberation Dialogues

5.1. *General Elements of Dialogues*

This section aims to formalize three types of dialogues: a backward deliberation dialogue, a forward deliberation dialogue and a mixed deliberation dialogue. Particularly, this section gives a formal definition of the intersection of backward and forward deliberation dialogues, as general as possible. In general, there are various ingredients associated with formal dialogues, e.g., locution, reply, commitment, turntaking, termination, and outcomes. In this paper, we only consider locutions, reply and outcomes because we think that they are necessary and sufficient factors to cover a key part of reconciliation typified by Kowalski and Toni's reconciliatory story. A general framework of reconciliatory dialogues equipped with all of the ingredients is beyond the scope of this paper although it is true that they make reconciliatory dialogues more realistic and sound. In our dialogue setting, unspecified number of players exchange moves during dialogues in which they always have their turn to put forward moves, they are not distinguished from a proponent and opponent, and their locutions are not subject to consistency check with their commitments.

Each move in dialogues consists of a *speech act* — the content of the move — and a *type* of dialogue — the context in which moves are put forward.

Definition 8. Let $a \in \mathcal{L}_0 \cup \mathcal{L}_1$. A *move* is a tuple $\langle speech\ act, type \rangle$ where $speech\ act \in \{claim(D\alpha), why(\alpha), since(\Phi \rightsquigarrow \alpha), fact(\alpha)\}$ and $type \in \{B, F\}$, i.e., backward deliberation dialogue or forward deliberation dialogue.

Let us define the set of allowed replies to a move.

Definition 9. Let M be a set of moves and $X \in \{B, F\}$. The following table depicts allowed replies to each locution.

Locutions	Replies
$\langle claim(D\alpha), X \rangle$	$\langle why(D\alpha), X \rangle$, $\langle claim(D\neg\alpha), X \rangle$
$\langle why(\alpha), X \rangle$	$\langle since(\Phi \rightsquigarrow \alpha), X \rangle, \langle fact(\alpha), X \rangle$
$\langle since(\Phi = \{\ldots, \beta, \ldots\} \rightsquigarrow \alpha), X \rangle$	$\langle why(\beta), X \rangle$
$\langle fact(\alpha), X \rangle$	

In the following, if $m \in M$ is a reply to $n \in M$, then we will say that m *attacks* n or $m \longrightarrow n$. Particularly, if $n = \langle claim(D\alpha), X \rangle$ and $m = \langle claim(D\neg\alpha), X \rangle$, then $m \longrightarrow n$ and $n \longrightarrow m$. No move attacks a set of moves nor another move with a different type.

A dialogue as a network of moves is defined using a dialogue framework. It is an abstract argumentation framework whose arguments and attacks are instantiated respectively by moves and attacks on the set of moves.

Definition 10. A dialogue framework is a pair $DF = \langle M, attacks \rangle$ where M is a set of moves and $attacks = \{\langle m, n \rangle | m, n \in M, m \longrightarrow n\}$.

Note that dialogue frameworks do not preserve the order in which agents put forward locutions, but preserve only the replying relation between moves. Agents

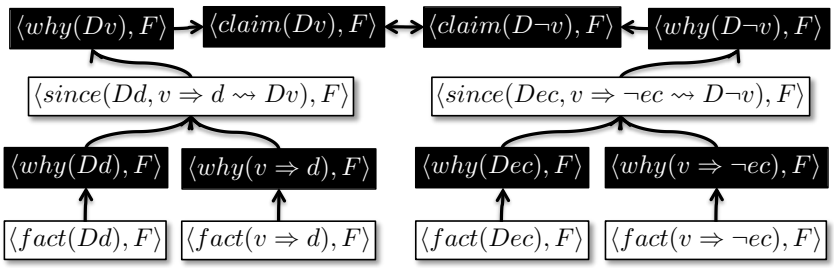

Figure 4. Forward deliberation dialogue with the overruled subject $\langle claim(Dv), F\rangle$ where nodes and links respectively represent moves and an attack relation.

dynamically construct a network of moves, i.e., a dialogue framework, by replying to a preceding move in it.

Dung's acceptability semantics evaluates acceptability of moves in dialogue frameworks. This is because it is rational to think that moves successfully replying to critical questions are worthy of acceptance. In general, dialogue frameworks are constructed by multiple agents who freely participate and make moves from their private knowledge bases. This knowledge is invisible to others. They can only see what they said during a dialogue.

Definition 11. Let DF be a dialogue framework. A *collaborative theory* built from DF, denoted by $T(DF)$, is the set $\{a \in \mathcal{L}_0 \cup \mathcal{L}_1 \mid$ There is a move m in DF whose speech act is $fact(a)\}$.

5.2. *Forward and Backward Deliberation Dialogues*

A forward deliberation protocol is defined using forward practical syllogisms.

Definition 12. Let M be a set of moves and $\mathcal{DF}$ be a set of dialogue frameworks and $Dh \in \mathcal{L}_0$. A *forward deliberation dialogue protocol* is a function $P_F : \mathcal{DF} \to 2^M$ where $m \in P_F(DF)$ if and only if the following hold.

- $m = \langle claim(Dh), F\rangle$ if $DF = \emptyset$.
- $m \notin DF$ and $\exists n \in DF$ such that $m \longrightarrow n$ if $DF \neq \emptyset$.

Moreover, if the speech act of m is $since(A \rightsquigarrow a)$, then $A \rightsquigarrow a$ is an application of forward practical syllogisms.

A dialogue framework DF is a *forward deliberation dialogue* iff DF is constructed by the forward deliberation protocol. The first move of DF is called the *subject* of the dialogue.

Example 1 (continued). Figure 4 presents an example of a forward deliberation dialogue with overruled subject $\langle claim(Dv), F\rangle$ in which white moves are justified and black ones are overruled. The collaborative theory is $T(DF) = \{Dd, v \Rightarrow d, Dec, v \Rightarrow \neg ec\}$.

A backward deliberation protocol is defined similarly.

Definition 13. Let M be a set of moves, $\mathcal{DF}$ be a set of dialogue frameworks and $Dh \in \mathcal{L}_0$. A *backward deliberation protocol* is a function $P_B : \mathcal{DF} \to 2^M$ where $m \in P_B(DF)$ if and only if the following hold.

- $m = \langle claim(Dh), B\rangle$ if $DF = \emptyset$.
- $m \notin DF$ and $\exists n \in DF$ such that $m \longrightarrow n$ if $DF \neq \emptyset$.

Moreover, if the speech act of m is $since(A \rightsquigarrow a)$, then $A \rightsquigarrow a$ is an application of

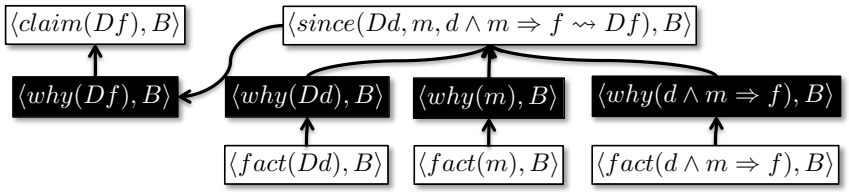

Figure 5. Backward deliberation dialogue with the justified subject $\langle claim(Df), B\rangle$.

backward practical syllogisms.

Definition 13 says that the backward deliberation protocol restricts a type of dialogue to a backward deliberation and an inference to backward practical syllogisms. We say that a dialogue framework DF is a *backward deliberation dialogue* iff DF is constructed by the backward deliberation protocol.

Example 1 (continued). Figure 5 presents an example of a backward deliberation dialogue in which subject $\langle claim(Df), B\rangle$ is justified. The collaborative theory is $T(DF) = \{Dd, m, d \wedge m \Rightarrow f\}$.

5.3. *Mixed Deliberation Dialogues*

We now define mixed deliberation dialogues by which agents search for and justify reconciliatory desires. We extend the forward and backward protocols so that claim moves in forward deliberation dialogues serve as a trigger to shift from backward to forward deliberation dialogues. Formally, a *mixed deliberation protocol* allows agents to put forward $\langle claim(D\alpha), B\rangle$ as a reply to $\langle why(D\alpha), F\rangle$ in mixed deliberation dialogues. It is defined based on the forward and backward deliberation dialogue protocols.

Definition 14. Let N be a set of moves and $\mathcal{DF}$ be a set of dialogue frameworks. A *mixed deliberation protocol* is a function $P_M : \mathcal{DF} \to 2^N$, where $m \in P_M(DF)$ iff $m \in P_F(DF) \cup P_B(DF)$ or, $m = \langle claim(Da), B\rangle$ and $\langle why(Da), F\rangle$ is in DF.

A dialogue framework DF is a *mixed deliberation dialogue* iff DF is constructed using the mixed deliberation protocol. Note that forward and backward deliberation dialogues are both special cases of mixed deliberation dialogues.

Example 1 (continued). Figure 6 presents an example of a mixed deliberation dialogue DF in which the subject $claim(Dh)$ is justified. The collaborative theory built from DF is $T(DF) = \{Dd, m, d \wedge m \Rightarrow f, r, h \wedge r \Rightarrow f\}$.

6. Correctness of Forward, Backward, and Mixed Deliberation Dialogues

6.1. *Forward Deliberation Dialogues for Justifying Subordinate Desires*

This section shows the relations between acceptability status of dialogue subjects and desires of three kinds (subordinate, superordinate, and reconciliatory desires) defined on collaborative theories. It gives dialogue agents rationale for accepting and agreeing to dialogue subjects.

We say a dialogue framework is *finite* if the number of moves in it is finite. We impose *closedness* on dialogue frameworks to associate status of subjects and subordinate, superordinate and reconciliatory desires. A dialogue framework is closed if it is finite and the corresponding dialogue protocol does not permit to put why

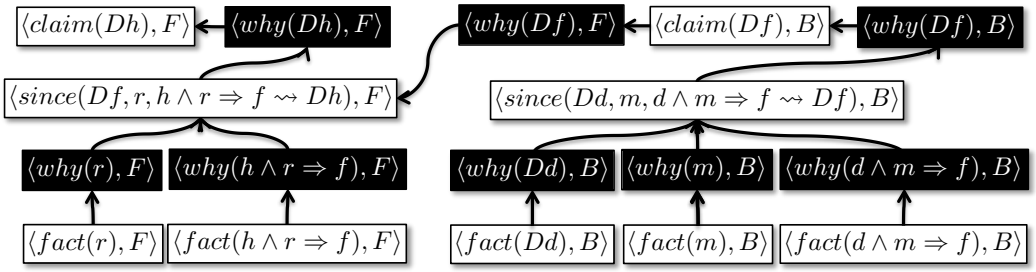

Figure 6. Mixed deliberation dialogue with the justified subject $\langle claim(Dh), F\rangle$.

and since moves forward.

Definition 15. Let DF be a finite dialogue framework and P be a protocol. DF is closed iff there is no $m \in P(DF)$ such that the speech act of m is $why(a)$, or it is $since(A \rightsquigarrow a)$ where, for all $c \in A$, there is a move n in DF such that the speech act of n is $since(B \rightsquigarrow b)$ and $c \in B$.

We show that subjects of forward deliberation dialogues interpret their dialogue goals, i.e., subordinate desires, defined on collaborative theories built through the dialogues. The following two lemmas guarantee that forward deliberation dialogues are *sound* in the sense that justified subjects are necessarily subordinate desires in the collaborative theory, and *complete* in the sense that subordinate desires in the collaborative theory are necessarily justified subjects.

Lemma 1. Let DF be a closed forward deliberation dialogue for which the subject is $\langle claim(Dh), F\rangle$. If $\langle claim(Dh), F\rangle$ is justified in DF, then there is $Dg \in T(DF)$ such that Dh is a subordinate desire of Dg in $T(DF)$.

The following Lemma 2 assures that forward deliberation dialogues are *complete* in the sense that subordinate desires in the collaborative theory are necessarily justified subjects.

Lemma 2. Let DF be a closed forward deliberation dialogue for which the subject is $\langle claim(Dh), F\rangle$. If there is $Dg \in T(DF)$ such that Dh is a subordinate desire of Dg in $T(DF)$ then $\langle claim(Dh), F\rangle$ is justified in DF.

Lemmas 1 and 2 imply Theorem 1.

Theorem 1. Let DF be a closed forward deliberation dialogue whose subject is $\langle claim(Dh), F\rangle$. $\langle claim(Dh), F\rangle$ is justified in DF iff there is $Dg \in T(DF)$ such that Dh is a subordinate desire of Dg in $T(DF)$.

6.2. *Backward Deliberation Dialogues for Justifying Superordinate Desires*

Subject status of backward deliberation dialogues interpret their dialogue goals, i.e., superordinate desires, defined on collaborative theories built through the dialogues. The following theorem can be shown similarly to Theorem 1.

Theorem 2. Let DF be a closed backward deliberation dialogue for which the subject is $\langle claim(Dh), B\rangle$. Actually, $\langle claim(Dh), B\rangle$ is justified in DF iff there is $Dg \in T(DF)$ such that Dh is a superordinate desire of Dg in $T(DF)$.

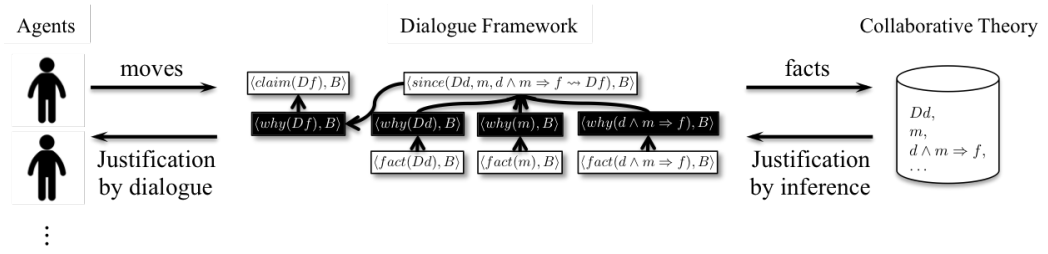

Figure 7. System overview of mixed deliberation dialogue

6.3. *Mixed Deliberation Dialogues for Justifying Reconciliatory Desires*

Subjects status of mixed deliberation dialogues interpret their dialogue goals, i.e., reconciliatory desires, defined on collaborative theories built through the dialogues. The following Lemma 3 guarantees that mixed deliberation dialogues are *sound* in the sense that justified subjects are necessarily reconciliatory desires in collaborative theories.

Lemma 3. Let $X \in \{B, F\}$ and DF be a closed mixed deliberation dialogue for which the subject is $\langle claim(Dh), X\rangle$. If $\langle claim(Dh), X\rangle$ is justified in DF then there is $Dg \in T(DF)$ such that Dh is a reconciliatory desire of Dg in $T(DF)$.

Lemma 4. Let $X \in \{F, B\}$ and DF be a closed mixed deliberation dialogue for which the subject is $\langle claim(Dh), X\rangle$. If $Dg \in T(DF)$ exists such that Dh is a reconciliatory desire of Dg in $T(DF)$, then $\langle claim(Dh), X\rangle$ is justified in DF.

Lemmas 3 and 4 imply Theorem 3.

Theorem 3. Let $X \in \{F, B\}$ and DF be a closed mixed deliberation dialogue for which the subject is $\langle claim(Dh), X\rangle$. There is $Dg \in T(DF)$ such that Dh is a reconciliatory desire of Dg in $T(DF)$ iff $\langle claim(Dh), X\rangle$ is justified in DF.

6.4. *Illustrative Example of Mixed Deliberation Dialogues*

This section gives an illustrative example to show how mixed deliberation dialogues unfold by agents. In Figure 7, we show information flows of our dialogue systems where moves from agents constitute a dialogue framework, and facts from the framework constitute a collaborative theory. Reconciliatory (resp. subordinate, superordinate) desires defined in the collaborative theory give a justification to mixed (resp. forward, backward) deliberation dialogues, and justified claims in the dialogue give a justification to multiagent decision making.

So far, however, we paid no attention to agent models (the leftmost component in Figure 7) such as agent's knowledge base nor agent's dialogue strategy. This is because we focus on dialogue protocols that should be distinguished from them. A dialogue protocol deals with moves agents are allowed to put forward in dialogues. On the other hand, a dialogue strategy deals with moves agents actually put forward in dialogues. Agents knowledge base defines moves she can make from her knowledge base, and this information affects her strategy of what to say in dialogues. However, it does not affect dialogue protocols of what moves she is allowed to put forward.

However, we think that agents' knowledge and strategies are necessary to show how a dialogue unfold by them. In this section, we assume two agents $agent_1$ and

$agent_2$ who have the following knowledge bases, $T_1, T_2 \subseteq \mathcal{L}_0 \cup \mathcal{L}_1$, respectively.

$$T_1 = \{Dd, v \Rightarrow d, m, r, a \Rightarrow ec\}$$
$$T_2 = \{Dec, v \Rightarrow ec, d \wedge m \Rightarrow f, h \wedge r \Rightarrow f\}$$

Moreover, we assume the simple strategy that each agent willingly and honestly participates in dialogues with no particular order. By permission of a given dialogue protocol, each agent puts forward $fact$ and *since* moves she can make from her knowledge base. She also puts forward *why* moves in any time and *claim* moves in the beginning of dialogues.

For example, consider the situation where $agent_1$ starts a dialogue on the subject whether v (i.e., "voting") is a subordinate (and therefore reconciliatory) desire or not.

Utterer	Move	Attack-to	Status
$agent_1$	$\langle claim(Dv), F\rangle$	-	overruled
$agent_2$	$\langle why(Dv), F\rangle$	$\langle claim(Dv), F\rangle$	overruled
$agent_1$	$\langle since(Dd, v \Rightarrow d \rightsquigarrow Dv), F\rangle$	$\langle why(Dv), F\rangle$	justified
$agent_2$	$\langle why(Dd), F\rangle$	$\langle since(Dd, v \Rightarrow d \rightsquigarrow Dv), F\rangle$	overruled
$agent_2$	$\langle why(v \Rightarrow d), F\rangle$	$\langle since(Dd, v \Rightarrow d \rightsquigarrow Dv), F\rangle$	overruled
$agent_1$	$\langle fact(Dd), F\rangle$	$\langle why(Dd), F\rangle$	justified
$agent_1$	$\langle fact(v \Rightarrow d), F\rangle$	$\langle why(v \Rightarrow d), F\rangle$	justified
$agent_2$	$\langle claim(D\neg v), F\rangle$	$\langle claim(Dv), F\rangle$	overruled
$agent_1$	$\langle why(D\neg v), F\rangle$	$\langle claim(D\neg v), F\rangle$	overruled
$agent_2$	$\langle since(Dec, v \Rightarrow \neg ec \rightsquigarrow D\neg v), F\rangle$	$\langle why(D\neg v), F\rangle$	justified
$agent_1$	$\langle why(Dec), F\rangle$	$\langle since(Dec, v \Rightarrow \neg ec \rightsquigarrow D\neg v), F\rangle$	overruled
$agent_1$	$\langle why(v \Rightarrow \neg ec), F\rangle$	$\langle since(Dec, v \Rightarrow \neg ec \rightsquigarrow D\neg v), F\rangle$	overruled
$agent_2$	$\langle fact(Dec), F\rangle$	$\langle why(Dec), F\rangle$	justified
$agent_2$	$\langle fact(v \Rightarrow \neg ec), F\rangle$	$\langle why(v \Rightarrow \neg ec), F\rangle$	justified

The above table shows a sequence of moves and their utterers, targets and status in the forward (and therefore mixed) deliberation dialogue where the subject $\langle claim(Dv), F\rangle$ is overruled. Now, consider the another situation where $agent_1$ starts another dialogue on the subject whether h (i.e., "head cmte administrative only") is a reconciliatory desire or not.

Utterer	Move	Attack-to	Status
$agent_2$	$\langle claim(Dh), F\rangle$	-	justified
$agent_1$	$\langle why(Dh), F\rangle$	$\langle claim(Dh), F\rangle$	overruled
$agent_2$	$\langle since(Df, r, h \wedge r \Rightarrow f \rightsquigarrow Dh), F\rangle$	$\langle why(Dh), F\rangle$	justified
$agent_1$	$\langle why(Df), F\rangle$	$\langle since(Df, r, h \wedge r \Rightarrow f \rightsquigarrow Dh), F\rangle$	overruled
$agent_1$	$\langle why(h \wedge r \Rightarrow f), F\rangle$	$\langle since(Df, r, h \wedge r \Rightarrow f \rightsquigarrow Dh), F\rangle$	overruled
$agent_1$	$\langle why(r), F\rangle$	$\langle since(Df, r, h \wedge r \Rightarrow f \rightsquigarrow Dh), F\rangle$	overruled
$agent_2$	$\langle claim(Df), B\rangle$	$\langle why(Df), F\rangle$	justified
$agent_2$	$\langle fact(h \wedge r \Rightarrow f), F\rangle$	$\langle why(h \wedge r \Rightarrow f), F\rangle$	justified
$agent_1$	$\langle fact(r), F\rangle$	$\langle why(r), F\rangle$	justified
$agent_1$	$\langle why(Df), B\rangle$	$\langle claim(Df), B\rangle$	overruled
$agent_2$	$\langle since(Dd, m, d \wedge m \Rightarrow f \rightsquigarrow Df), B\rangle$	$\langle why(Df), B\rangle$	justified
$agent_1$	$\langle why(Dd), B\rangle$	$\langle since(Dd, m, d \wedge m \Rightarrow f \rightsquigarrow Df), B\rangle$	overruled
$agent_1$	$\langle why(m), B\rangle$	$\langle since(Dd, m, d \wedge m \Rightarrow f \rightsquigarrow Df), B\rangle$	overruled
$agent_1$	$\langle why(d \wedge m \Rightarrow f), B\rangle$	$\langle since(Dd, m, d \wedge m \Rightarrow f \rightsquigarrow Df), B\rangle$	overruled
$agent_1$	$\langle fact(Dd), B\rangle$	$\langle why(Dd), B\rangle$	justified
$agent_1$	$\langle fact(m), B\rangle$	$\langle why(m), B\rangle$	justified
$agent_2$	$\langle fact(d \wedge m \Rightarrow f), B\rangle$	$\langle why(d \wedge m \Rightarrow f), B\rangle$	justified

The above table shows a sequence of moves and their utterers, targets and status in the mixed deliberation dialogue where the subject $\langle claim(Dh), F\rangle$ is justified.

We can also see that the subject $\langle claim(Da), F\rangle$ is justified in the following

forward (and therefore mixed deliberation) dialogue started by $agent_2$. Note that each reconciliatory desire justified in dialogues is derived neither from T_1 nor T_2 using forward and backward practical syllogisms.

Utterer	Move	Attack-to	Status
$agent_2$	$\langle claim(Da), F\rangle$	-	justified
$agent_1$	$\langle why(Da), F\rangle$	$\langle claim(Da), F\rangle$	overruled
$agent_1$	$\langle since(Dec, a \Rightarrow ec \rightsquigarrow Da), F\rangle$	$\langle why(Da), F\rangle$	justified
$agent_2$	$\langle why(Dec), F\rangle$	$\langle since(Dec, a \Rightarrow ec \rightsquigarrow Da), F\rangle$	overruled
$agent_2$	$\langle why(a \Rightarrow ec), F\rangle$	$\langle since(Dec, a \Rightarrow ec \rightsquigarrow Da), F\rangle$	overruled
$agent_2$	$\langle fact(Dec), F\rangle$	$\langle why(Dec), F\rangle$	justified
$agent_1$	$\langle fact(a \Rightarrow ec), F\rangle$	$\langle why(a \Rightarrow ec), F\rangle$	justified

7. Related Work and Discussions

Focusing on superordinate or underlying desires furthers reconciliation. Fisher et al. (Fisher et al. 1992) argue that paying attention to stakeholders' interests helps to find reconciliation in negotiation. Brett (Brett 2014) says that negotiation theory distinguishes a position and an interest where a proposition is what negotiators say they want and an interest is the needs of concerns that underlie positions.

In argument-based negotiation and deliberation, the research studies (Amgoud et al. 2009, Hulstijn and van der Torre 2004, Modgil and Luck 2009) use practical reasoning to deal with desire derivation and generation mechanisms for knowledge-dependent and context-dependent desires. They, however, do not address the situations in which there is no means to achieve original nor derivative desires. Rahwan et al. (Rahwan et al. 2007) argue that underlying goals improve negotiation processes and consider desires hierarchized in advance. In contrast to their approach, we assume the situation where desires are structurized as a result of inference by forward and backward practical syllogisms. Our approach is necessary when agents have incomplete desires, as well as incomplete knowledge. Hitchcock et al. (Hitchcock et al. 2001) and McBurney et al. (McBurney et al. 2007) propose deliberation dialogue frameworks, e.g., DDF (McBurney et al. 2007), equipped with fundamental elements for deliberation dialogues such as locutions, commitments, and termination. Kok et al. (Kok et al. 2010) give an argumentation framework for deliberation dialogue taking into account agent's preference. However, these frameworks do not address evaluation of their correctness. In this paper, we gave proof-based evaluation for correctness of our dialogues in terms of dialogue goals. Fan and Toni (Fan and Toni 2012) use an assumption-based argumentation framework to relate successful dialogues with admissible arguments in the framework. Wells and Reed (Wells and Reed 2006) handle a shift from persuasion to negotiation, and demonstrate its effectiveness using an example. Although we also use argumentation frameworks and dialectical shifts to define our dialogues, the type of dialogue and the goal of dialogue, we focus on, are uniquely deliberation and reconciliatory, respectively.

Kido and Ohsawa (Kido and Ohsawa 2013) propose a reconciliatory argumentation system instantiating Dung's argumentation framework with modal propositional language and rules of practical inference. Our dialogue-based reconciliation has an advantage over argumentation-based in the sense that it handles circumstances under which individual agents cannot make a reconciliatory argument by themselves because of lack of knowledge, but can make such an argument for mutual coverage. Moreover, dialogue-based approaches succeed in capturing dynamic aspects of interaction where individual agents build a theory collaboratively during a dialogue. On the other hand, this paper has the limitation that dialogue proto-

cols allow agents to use only practical syllogisms and allow them to attack only moves except facts. We think that our work can be extended to a general reconciliatory dialogue by utilizing the persuasion dialogues mentioned above and inquiry dialogues (Black and Hunter 2009) allowing agents to share knowledge to jointly construct arguments or dialectical trees.

Much work for formal deliberation dialogues has been inspired by studies of formal persuasion dialogues. Prakken (Prakken 2005) provides a formal dialogue system and shows that, under some conditions, a proponent wins in a dialogue iff a topic of the dialogue is defeasibly derived from agreed information. We also adopt a similar manner in giving the correctness of our dialogue protocols because we agree with Carlson's idea (Carlson 1983), cited by Prakken (Prakken 2005, 2006), that "whereas logic defines the conditions under which a proposition is true, dialogue systems define the conditions under which an utterance is appropriate, and this is the case if the utterance furthers the goal of the dialogue in which it is made." However, our dialogue protocols are completely different to Prakken's because of differences of dialogue types and goals. It is true that there are various ingredients associated with formal dialogues, e.g., locution, reply, commitment, turntaking, termination, and outcomes, and our protocols deal with some of them, in contrast to Prakken's protocols. However, our approach does not conflict with his approach because, as Prakken mentioned in (Prakken 2005), his structure of dialogues is especially suited for "verbal struggles" and he does not claim that all dialogues should or do conform to the structure. In fact, our dialogues have little interest in commitment because consistency evaluation of locutions in terms of commitment does not directly contribute to the solution of Kowalski and Tonis scenario. This is supported by the fact that Kowalski and Toni's scenario makes no mention of agent's commitment in the process of reaching the reconciliatory solution. Our dialogues neither have an interest in an wining-or-losing criterion because agents do not compete against each other with the different roles, a proponent and an opponent, but collaborate each other to find a reconciliation. Turntaking is neither an essential because agents having the same role do not need to be distinguished. A termination criterion is not essential because Dung's grounded semantics allows us to give outcomes of any particular situation of dialogues. It is true that all of these factors make our dialogues more realistic and sound. However, pursuing a generality of reconciliatory dialogues is beyond the scope of this paper.

8. Conclusions and Future Work

We proposed a mixed deliberation dialogue for reconciliatory desires. A mixed deliberation dialogue is defined as a combination of forward and backward deliberation dialogues for which the goals are subordinate and superordinate desires, respectively. We showed the correctness of dialogues based on the fact that the subject of closed mixed deliberation dialogue is justified iff the subject is a reconciliatory desire of a desire in the collaborative theory, i.e., iff the dialogue satisfies its goal. Weaknesses of our formalization are, first, that it restricts inference mechanisms to only forward and backward practical syllogisms, and second, that it does not allow agents to challenge or rebut facts put forward in dialogues. We will utilize existing persuasion and inquiry dialogues to address these issues.

Acknowledgement

This research was partially supported by the Telecommunication Advancement Foundation, Japan.

References

L. Amgoud, C. Devred, and M. Lagasquie-Schiex. A constrained argumentation system for practical reasoning. In *Proc. of the Fifth International Workshop on Argumentation in Multi-Agent Systems*, pages 37–56, 2009.

T. J. M. Bench-Capon and H. Prakken. Justifying actions by accruing arguments. In *Proc. of the 1st International Conference on Computational Models of Argument*, pages 247–258, 2006.

E. Black and A. Hunter. An inquiry dialogue system. *JAAMAS*, 19:173–209, 2009.

J. M. Brett. *Negotiating Globally: How to Negotiate Deals, Resolve Disputes, and Make Decisions Across Cultural Boundaries (Jossey-Bass Business & Management)*. Jossey-Bass, 3rd Edition, 2014.

L. Carlson. *Dialogue Games: an Approach to Discourse Analysis*. Reidel Publishing Company, Dordrecht, 1983.

P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

X. Fan and F. Toni. Argumentation dialogues for two-agent conflict resolution. In *Proc. of the 4th International Conference on Computational Models of Argument*, pages 249–260, 2012.

R. Fisher, W. Ury, and B. Patton. *Getting to Yes: Negotiating Agreement Without Giving In*. Houghton Mifflin Harcourt, 2nd Edition, 1992.

C. L. Hamblin. *Fallacies*. Methuen, 1970.

J. Hintikka. *Language-Games for Quantifiers*, pages 46–72. Oxford: Blackwell, in american philosohpical quarterly monograph series 2: studies in logical theory edition, 1968.

D. Hitchcock, P. McBurney, and S. Parsons. A framework for deliberation dialogues. In *Proc. of the 4th biennial conference of the Ontario Society for the Study of argumentation*, pages 1–24, 2001.

J. Hulstijn and L. van der Torre. Combining goal generation and planning in an argumentation framework. In *Proc. of the 10th International Workshop on Non-Monotonic Reasoning*, pages 212–218, 2004.

H. Kido and F. Cerutti. Shift from forward to backward deliberation in serach of reconciliation. In *Proc. of the 13th Pacific Rim International Conference on Artificial Intelligence (PRICAI 2014)*, pages 920–928, 2014.

H. Kido and Y. Ohsawa. Justifying underlying desires for argument-based reconciliation. In *Proc. of the 2nd International Workshop on Theory and Applications of Formal Argumentation (TAFA 2013)*, pages 143–157, 2013.

E. M. Kok, J. J. C. Meyer, H. Prakken, and G. A. W. Vreeswijk. A formal argumentation framework for deliberation dialogues. In *Proc. of the 7th International Workshop on Argumentation in Multi-Agent Systems*, pages 31–48, 2010.

R. A. Kowalski and F. Toni. Argument and reconciliation. In *Proc. of the 5th Generation Computer Systems Workshop on Application of Logic Programming to Legal Reasoning*, pages 9–16, 1994.

P. Lorenzen. *Ein dialogisches konstruktivittskriterium*, pages 193–200. Oxford: Pergamon, in infinitistic methods edition, 1961.

P. McBurney, D. Hitchcock, and S. Parsons. The eightfold way of deliberation dialogue. *International Journal of Intelligent Systems*, 22(1):95–132, 2007.

S. Modgil and M. Luck. Argumentation based resolution of conflicts between desires and normative goals. In *Proc. of the 5th International Workshop on Argumentation in Multi-Agent Systems (ArgMAS 2009)*, pages 19–36, 2009.

H. Prakken. Coherence and flexibility in dialogue games for argumentation. *Journal of Logic and Computation*, 15:1009–1040, 2005.

H. Prakken. Formal systems for persuasion dialogue. *The Knowledge Engineering Review*, 21(2):163–188, 2006.

I. Rahwan, P. Pasquier, L. Sonenberg, and F. Dignum. On the benefits of exploiting underlying goals in argument-based negotiation. In *Proc. of the 22nd National Conference on Artificial Intelligence (AAAI 2007)*, pages 116–121, 2007.

F. H. van Eemeren, R. Grootendorst, and F. S. Henkemans. *Fundamentals of Argumentation Theory: A Handbook of Historical Backgrounds and Contemporary Developments*. Routledge, 1996.

S. Wells and C. Reed. Knowing when to bargain. In *Proc. of the 1st International Conference on Computational Models of Argument*, pages 235–246, 2006.

Appendix A. Proofs

Proposition 1. Let $Dg, Dh \in \mathcal{L}_0$. Dh is a superordinate desire of Dg in T iff there is $\Sigma_1 \subseteq T$ such that $\Sigma_1 \cup \{Dg\} \vdash_{BPS} Dh$ and there is no $\Sigma_2 \subseteq T$ such that $\Sigma_2 \cup \{Dt\} \vdash_{BPS} D\neg h$, for all $Dt \in T \cup \{Dg\}$.

Proof. It is obvious that, for any $\Sigma \subseteq T$ and $Dx, Dy \in \mathcal{L}_0$, $\Sigma \cup \{Dx\} \vdash_{FPS} Dy$ iff $\Sigma \cup \{Dy\} \vdash_{BPS} Dx$. So, Dh is a superordinate desire of Dg in T iff (1) there is $\Sigma_1 \subseteq T$ such that $\Sigma_1 \cup \{Dg\} \vdash_{BPS} Dh$ and (2) there is no $\Sigma_2 \subseteq T$ such that $\Sigma_2 \cup \{Dt\} \vdash_{BPS} D\neg h$, for all $Dt \in T \cup \{Dg\}$. □

Lemma 1. Let DF be a closed forward deliberation dialogue for which the subject is $\langle claim(Dh), F\rangle$. If $\langle claim(Dh), F\rangle$ is justified in DF, then there is $Dg \in T(DF)$ such that Dh is a subordinate desire of Dg in $T(DF)$.

Proof. The following proof still holds when $a_{i+1} \wedge b_i \Rightarrow a_i$ is replaced by $\neg a_{i+1} \wedge b_i \Rightarrow \neg a_i$. Since DF is closed and $\langle claim(Dh), F\rangle$ is justified, $fact(Dh)$ appears in DF or there is $Da_1 \in \mathcal{L}_0 \cup \mathcal{L}_1$ such that $fact(Da_1)$ appears in DF and, for all $i(1 \leq i \leq n)$, there are $a_i, b_i \in \mathcal{L}_0 \cup \mathcal{L}_1$ such that $fact(b_i)$, $fact(a_{i+1} \wedge b_i \Rightarrow a_i)$ and $since(Da_i, b_i, a_{i+1} \wedge b_i \Rightarrow a_i \rightsquigarrow Da_{i+1})$ appear in DF where $a_{n+1} = h$. Therefore, $Da_1 \in T(DF)$ and $b_i, a_{i+1} \wedge b_i \Rightarrow a_i \in T(DF)$, for all $i(1 \leq i \leq n)$.

Based on proof by contradiction, we show there is no $\Sigma \subseteq T(DF)$ such that $\Sigma \cup \{Dt\} \vdash_{FPS} D\neg h$, for all $Dt \in T(DF) \cup \{Dg\}$. If $\Sigma \cup \{Dt\} \vdash_{FPS} D\neg h$ then $fact(Da_1)$, $fact(b_i)$ and $fact(a_{i+1} \wedge b_i \Rightarrow a_i)$ appear in DF, for all $i(1 \leq i \leq n)$, where $a_{n+1} = \neg h$. So, $\langle why(Da_i), F\rangle$, $\langle why(b_i), F\rangle$ and $\langle why(a_{i+1} \wedge b_i \Rightarrow a_i), F\rangle$ are all attacked, for all $i(1 \leq i \leq n)$. Therefore, it is not the case that $\langle claim(D\neg h), F\rangle$ is overruled. Since $\langle claim(D\neg h), F\rangle$ attacks $\langle claim(Dh), F\rangle$, $\langle claim(Dh), F\rangle$ cannot be justified in DF. □

Lemma 2. Let DF be a closed forward deliberation dialogue for which the subject is $\langle claim(Dh), F\rangle$. If there is $Dg \in T(DF)$ such that Dh is a subordinate desire of Dg in $T(DF)$ then $\langle claim(Dh), F\rangle$ is justified in DF.

Proof. The following proof still holds when $a_{i+1} \wedge b_i \Rightarrow a_i$ is replaced by $\neg a_{i+1} \wedge b_i \Rightarrow \neg a_i$. There is $\Sigma_1 \subseteq T(DF)$ such that $\Sigma_1 \cup \{Dg\} \vdash_{FPS} Dh$ and there is no $\Sigma_2 \subseteq T(DF)$ such that $\Sigma_2 \cup \{Dt\} \vdash_{FPS} D\neg h$, for all $Dt \in T(DF) \cup \{Dg\}$. Since $\langle fact(Dg), F\rangle$ appears in DF and there are $a_i, b_i \in \mathcal{L}_0 \cup \mathcal{L}_1$ such that $\langle fact(b_i), F\rangle$ and $\langle fact(a_{i+1} \wedge b_i \Rightarrow a_i), F\rangle$ appear in DF, for all $i(1 \leq i \leq n)$, where $a_1 = g$ and $a_{n+1} = h$. So, for all $i(1 \leq i \leq n)$, $\langle why(Da_i), F\rangle$, $\langle why(b_i), F\rangle$ and $\langle why(a_{i+1} \wedge b_i \Rightarrow a_i), F\rangle$ are all attacked.

Since there is $i(1 \leq i \leq n)$ such that no moves $\langle fact(Dc_1), F\rangle$, $\langle fact(d_i), F\rangle$ or $\langle fact(c_{i+1} \wedge d_i \Rightarrow c_i), F\rangle$ appear in DF where $c_{n+1} = \neg h$, there is $i(1 \leq i \leq n)$ such that $\langle why(Dc_i), F\rangle$, $\langle why(d_i), F\rangle$ or $\langle why(c_{i+1} \wedge d_i \Rightarrow c_i), F\rangle$ is not attacked and justified. Since $\langle claim(D\neg h), F\rangle$ is overruled, $\langle claim(Dh), F\rangle$ is justified. □

Theorem 1. Let DF be a closed forward deliberation dialogue whose subject is $\langle claim(Dh), F\rangle$. $\langle claim(Dh), F\rangle$ is justified in DF iff there is $Dg \in T(DF)$ such that Dh is a subordinate desire of Dg in $T(DF)$.

Proof. Immediate from Lemmas 1 and 2. □

Theorem 2. Let DF be a closed backward deliberation dialogue for which the subject is $\langle claim(Dh), B\rangle$. Actually, $\langle claim(Dh), B\rangle$ is justified in DF iff there is $Dg \in T(DF)$ such that Dh is a superordinate desire of Dg in $T(DF)$.

Proof. As a result of Proposition 1, the proof is same as Theorem 1. □

Lemma 3. Let $X \in \{B, F\}$ and DF be a closed mixed deliberation dialogue for which the subject is $\langle claim(Dh), X\rangle$. If $\langle claim(Dh), X\rangle$ is justified in DF then there is $Dg \in T(DF)$ such that Dh is a reconciliatory desire of Dg in $T(DF)$.

Proof. It is sufficient to consider a mixed deliberation dialogue DF whose subject $claim(Dh)$ is not justified in neither only forward nor backward deliberation dialogues appeared in DF since otherwise the proof is reduced to Theorems 1 and 2. So, the only situation is that the dialogue starts with a forward deliberation dialogue whose subject $\langle claim(Dh), F\rangle$ is not justified in the dialogue, and then, shifts to backward deliberation dialogues and one, denoted by DF_b, of them makes it justified in DF. The subject, denoted by $\langle claim(Dg), B\rangle$, of DF_b is justified because otherwise DF_d cannot make $\langle claim(Dh), F\rangle$ justified in DF. From Theorem 1, this means that Dg is a superordinate desire of a desire in $T(DF)$. Since $\langle claim(Dh), F\rangle$ is justified in DF by the existence of $claim(Dg)$, Theorem 2 implies that Dh is a subordinate desire of Dg in $T(DF)$. □

Lemma 4. Let $X \in \{F, B\}$ and DF be a closed mixed deliberation dialogue for which the subject is $\langle claim(Dh), X\rangle$. If $Dg \in T(DF)$ exists such that Dh is a reconciliatory desire of Dg in $T(DF)$, then $\langle claim(Dh), X\rangle$ is justified in DF.

Proof. It is sufficient to consider Dh that is neither subordinate nor superordinate desires of Dg because otherwise these cases reduce the proof to Theorems 1 and 2. Let $Di \in \mathcal{L}_0 \setminus T(DF)$ be a desire such that Dh is a subordinate desire of Di and Di is a superordinate desire of Dg in $T(DF)$. From Theorem 1, there is a backward deliberation dialogue, denoted by BD, whose subject $\langle claim(Di), B\rangle$ is justified in the dialogue. On the other hand, we can consider the forward deliberation dialogue, denoted by FD, whose root is $\langle claim(Dh), F\rangle$ where the root is overruled due to the fact that $\langle why(Di), F\rangle$ is justified. However, $\langle claim(Dh), F\rangle$ becomes justified in the mixture of BD and FD because $\langle claim(Di), B\rangle$ makes $\langle why(Di), F\rangle$ overruled. □

Theorem 3. Let $X \in \{F, B\}$ and DF be a closed mixed deliberation dialogue for which the subject is $\langle claim(Dh), X\rangle$. There is $Dg \in T(DF)$ such that Dh is a reconciliatory desire of Dg in $T(DF)$ iff $\langle claim(Dh), X\rangle$ is justified in DF.

Proof. Immediate from Lemmas 3 and 4. □