ECAI 2012
Luc De Raedt et al. (Eds.)

doi:10.3233/978-1-61499-098-7-756

An Empirical Study of Argumentation Schemes for Deliberative Dialogue

Alice Toniolo[1] and **Timothy J. Norman**[1] and **Katia Sycara**[2]

Abstract. Collaborative decision making among agents in a team is a complex activity, and tasks to achieve individual objectives may conflict in a team context. A number of argumentation-based models have been proposed to address the problem, the rationale being that the revelation of background information and constraints can aid in the discovery and resolution of conflicts. To date, however, no empirical studies have been conducted to substantiate these claims. In this paper, we discuss a model, grounded on argumentation schemes, that captures potential conflicts due to scheduling and causality constraints, and individual goals and norms. We evaluate this model in complex collaborative planning problems and show that such a model facilitates the sharing of relevant information pertaining to plan, goal and normative conflicts. Further, we show that this focussed information sharing leads to more effective conflict resolution, particularly in the most challenging problems.

1 Introduction

If autonomous agents are able to be effective in fulfilling their objectives while complying with societal norms in a team context, they must coordinate their activities. This is critical in domains in which the plans of individuals are highly interdependent. Decisions made by individual agents may interfere with those of others due to, for example, concurrent resource use or different normative expectations about how things are done. How can we develop effective mechanisms for agents to establish agreements on how to act together?

Argumentation-based models of dialogue enable agents to effectively resolve conflicts of opinion. This has been shown by empirical studies in different types of dialogue; for example, in persuasion dialogue, using argumentation an agent is able to influence another's intentions [4], and in negotiation, argumentation enables more efficient reallocation of resources [2]. Earlier research on the use of argumentation-based models in multi-agent systems addressed the problem of creating consistent shared plans [6]. Recently, Atkinson and Bench-Capon [1] proposed a model of deliberative dialogue based on argumentation schemes and focussed on constructing a common plan for a joint goal. Toniolo et al. [7] presented a formalisation of argumentation schemes appropriate for deliberative dialogue where conflicts may also arise due to differing objectives and normative constraints. However, existing research has focussed on demonstrating *that* argumentation-based models are useful through the use of extended examples. No rigorous assessment of *how* the information conveyed using arguments affects the resolution of conflicts has been performed. In this paper we consider complex collaborative planning problems where agents must resolve conflicts among interdependent plans. At the same time, each agent aims to satisfy its individual goals and normative constraints. We integrate a model of argumentation schemes with a rigorous argumentation framework to be used by agents to resolve conflicts between interdependent plans. We empirically evaluate this model and show that the use of this model of argumentation schemes facilitates the identification of conflicts that hamper the execution of collaborative tasks. We then show that the information about plan, norm and goal conflicts shared during the dialogue leads to more effective conflict resolution.

2 Argumentation System

In this work we consider agents with individual plans to achieve their own objectives. Each agent's plan is consistent with respect to its knowledge, and compliant with local norms and resource constraints. Agents, then, discuss interdependent actions, norms and goals to resolve conflicts between their plans. The model evaluated in this paper uses argumentation schemes for practical reasoning [1, 7]. These schemes enable agents to identify conflicts among interdependent plans. The model supports the identification of scheduling and causal inconsistencies of actions, norm violation and incompatible goals.

Here we describe a dialogue system that integrates the argumentation schemes employed into an instantiation of the argumentation framework for deliberation proposed by Kok et al. [3]. An argumentation system includes a language for discussing agents' plans, a model of arguments, a set of defeasible relations among arguments, and a dialogue protocol. It is defined as a pair $\langle \mathcal{L}, \mathcal{D} \rangle$ where $\mathcal{L}$ is a logic for defeasible argumentation and $\mathcal{D}$ is the dialogue system.

2.1 Argumentation logic

A tuple $\mathcal{L} = \langle \mathcal{L}_t, \mathcal{R}_{in}, Args, \mathcal{R}_{def}, \mathcal{R}_{sup} \rangle$ defines the argumentation logic, where $\mathcal{L}_t$ is the topic language, $\mathcal{R}_{in}$ is a set of inference rules, $Args$ is a set of arguments, $\mathcal{R}_{def}$ and $\mathcal{R}_{sup}$ represent sets of defeat and support relations among arguments respectively. The topic language for arguments $\mathcal{L}_t$ is based on situation calculus and the inference rules $\mathcal{R}_{in}$ are defined by the argumentation schemes. The arguments include sub-sets $Args = Args_c \cup Args_p \cup Args_n \cup Args_g$ that deal with concurrent actions, causal links between actions, norms and goals respectively. The defeat $\mathcal{R}_{def}$ and support $\mathcal{R}_{sup}$ relations are expressed by the critical questions *CQ1,CQ2,CQ3,CQ4*.

Agents' plans. The formal representation of plans for agents is based on situation calculus, a well-known language for modelling dynamic domains, extended for representing time and norms [5]. Here we outline the plan elements relevant to the argumentation framework. The set of agents is Agt where $x, y \in Agt$, the set of actions is $\mathcal{A} = \{A_1, \ldots, A_l\}$, and $\mathcal{T} = \{T_1, \ldots, T_q\}$ is the set of time stamps. R represents a feature of the world that holds in a situation.

---

[1] Department of Computing Science, University of Aberdeen, Scotland, UK
[2] Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, US

Intuitively, the evolution of the world is represented as a tree where the root is the initial situation, nodes are possible world situations and arcs are actions that modify situations. This tree identifies the planning domain D^x for an agent x. The model includes norms that define what an agent is obliged to do or forbidden from doing in terms of actions or features of the world under certain conditions. Active norms branch nodes or arcs of the situation tree. A plan for an agent x is identified by a path from the root node to a node in which an overall goal Ψ is achieved, defined by sentences $\Psi \equiv \psi_1 \wedge \cdots \wedge \psi_m$. This path is a sequence of situations, each of which is scheduled on a common time line by a function that indicates its starting time.

Definition 1 *A temporally grounded plan* $\vec{P}^x$ *for agent* x *is defined as a solution to a planning problem identified by a domain* D^x *and a goal* Ψ. $\vec{P}^x = \langle [T_1]A_1; \ldots; [T_n]A_n \rangle$ *is a sequence of actions where* $[T_i]A_i$ *indicates an action* A_i *responsible for the transition from one situation to the next situation starting at time* T_i*.*

A durative action is represented with two instantaneous actions that identify its beginning and end; i.e. a durative action A_{dk}, where $[T_j]A_j = [T_{k1}]\mathit{begin}(A_{dk})$ and $[T_h]A_h = [T_{k2}]\mathit{end}(A_{dk})$, is part of the plan $\vec{P}^x$ if both actions $[T_j]A_j$ and $[T_h]A_h$ are included in $\vec{P}^x$.

Given a fully specified individual plan $\vec{P}^x$ for agent x, the following elements are required for the dialogue. A sequence of actions to achieve a goal ψ_k is $\vec{P}^x(\psi_k)$. Each action A_{dk} in the plan has preconditions $\mathcal{P}_{A_{dk}} = \{R_a, \ldots, R_u\}$ that hold in a situation before $[T_{k1}]\mathit{begin}(A_{dk})$ and effects $\mathcal{E}_{A_{dk}} = \{R_b, \ldots, R_w\}$ which hold in the situation following $[T_{k2}]\mathit{end}(A_{dk})$. An action can have effects that provide preconditions for subsequent actions. These relations are called causal links, $\mathit{cLink}(R, A_{dh}, A_{dk})$, where R is a precondition of A_{dk} provided as an effect of A_{dh}. A temporal relation $\mathit{occOver}_{\mathcal{T}}(A_{dk}, T_1, T_2)$ indicates that the execution of A_{dk} overlaps the interval of time between T_1 and T_2. Norms can forbid F or oblige O a state of the world to hold or an action to be performed (where $O \equiv F\neg$). The activation of the norm is induced by premises $\mathcal{N}_{Prem} = \{R_d, \ldots, R_z\}$ and an action A_{dp}. The following plan sentences ℓ_{ti} within $\mathcal{L}_t$ are used by agents during dialogue:

Definition 2 *Given* x*'s plan* $\vec{P}^x$*:* $\mathit{Perform}(x, A_{dk}, T_{k1}, T_{k2})$ *indicates that* x *performs action* A_{dk} *delimitated by* $[T_{k1}]\mathit{begin}(A_{dk})$, $[T_{k2}]\mathit{end}(A_{dk})$*;* $\mathit{Hold}(x, R, T_k)$ *indicates that for agent* x *a feature* R *holds at* T_k*;* $\mathit{Hold}(x, \{R_1, \ldots, R_n\}, T_k)$ *is a set of features holding at* T_k*;* $\mathit{Achieve}(x, \psi_k)$ *indicates that* x *achieves goal* ψ_k*.*

Arguments. The formal structure of the arguments is based on argumentation schemes for deliberative dialogue [7]. A scheme includes the argument and critical questions that may be used to construct new arguments. We describe here the claim and four schemes for norm, action and goal conflicts between individual plans. We assume that an agent x, the proponent, informs agent y, the opponent, about an action A_{dk} that it intends to perform. If the agents agree, A_{dk} will be included in the shared set of committed actions $\vec{P}^{xy}$.

Definition 3 (*Arg*$_I$-Claim) *The claim is an argumentation scheme used by an agent* x *to describe a proposed action* A_{dk} *with preconditions* $\mathcal{P}_{A_{dk}}$*, effects* $\mathcal{E}_{A_{dk}}$ *and the goal* ψ_k *that* A_{dk} *contributes to.*

- Given preconditions $\mathit{Hold}(x, \mathcal{P}_{A_{dk}}, T_{k1})$
- $\mathit{Perform}(x, A_{dk}, T_{k1}, T_{k2})$
- will bring about $\mathit{Hold}(x, \mathcal{E}_{A_{dk}}, T_{k2})$
- and will contribute to $\mathit{Achieve}(x, \psi_k)$ [where $A_{dk} \in \vec{P}^x(\psi_k)$],

$\Rightarrow$ then $\mathit{Perform}(x, A_{dk}, T_{k1}, T_{k2})$

Definition 4 (*Arg*$_c$-Concurrent Actions) *Argument* Arg_c *is used for potential conflicts among preconditions and effects of a proposed action* A_{dk} *and a concurrent action* A_{dh} *in the plan of* y*.*

- Given preconditions $\mathit{Hold}(x, \mathcal{P}_{A_{dk}}, T_{k1})$,
- $\mathit{Perform}(x, A_{dk}, T_{k1}, T_{k2})$,
- will bring about $\mathit{Hold}(x, \mathcal{E}_{A_{dk}}, T_{k2})$ [$R \in \{\mathcal{P}_{A_{dk}} \cup \mathcal{E}_{A_{dk}}\}$],
- and given preconditions $\mathit{Hold}(y, \mathcal{P}_{A_{dh}}, T_{h1})$,
- $\neg\mathit{Perform}(y, A_{dh}, T_{h1}, T_{h2})$,
- will bring about $\neg\mathit{Hold}(y, \mathcal{E}_{A_{dh}}, T_{h2})$ [$\neg R \in \{\mathcal{P}_{A_{dh}} \cup \mathcal{E}_{A_{dh}}\}$]
- and will contribute $\neg\mathit{Achieve}(y, \psi_h)$, [$A_{dh} \in \vec{P}^y(\psi_h)$]
- and $\mathit{occOver}_{\mathcal{T}}(A_{dk}, T_{h1}, T_{h2})$

$\Rightarrow$ then $\neg\mathit{Perform}(x, A_{dk}, T_{k1}, T_{k2}) \wedge$
$\mathit{Perform}(y, A_{dh}, T_{h1}, T_{h2})$.

Two concurrent actions A_{dh} and A_{dk} can be executed only if their preconditions hold and the effects of one action do not negate the effects of the other. The effects of A_{dh} should also not negate the preconditions of A_{dk} and vice-versa. When these conditions are violated, action A_{dk} cannot be adopted in the common plan $\vec{P}^{xy}$ and opponent y may use argument Arg_c to explain this conflict.

Definition 5 (*Arg*$_p$-Plan Constraints) *An argument* Arg_p *expresses a possible threat that a proposed action* A_{dk} *may represent to causal links between actions in the opponent's plan.*

- Given preconditions $\mathit{Hold}(x, \mathcal{P}_{A_{dk}}, T_{k1})$,
- $\mathit{Perform}(x, A_{dk}, T_{k1}, T_{k2})$,
- will bring about $\mathit{Hold}(x, \mathcal{E}_{A_{dk}}, T_{k2})$ [$R \in \{\mathcal{P}_{A_{dk}} \cup \mathcal{E}_{A_{dk}}\}$],
- Given $\mathit{cLink}(\neg R, A_{da}, A_{db})$
- $\neg\mathit{Perform}(y, A_{da}, T_{a1}, T_{a2})$ and $\neg\mathit{Perform}(y, A_{db}, T_{b1}, T_{b2})$,
- and $\mathit{occOver}_{\mathcal{T}}(A_{dk}, T_{a1}, T_{b2})$

$\Rightarrow$ then $\neg\mathit{Perform}(x, A_{dk}, T_{k1}, T_{k2}) \wedge$
$\mathit{Perform}(y, A_{da}, T_{a1}, T_{a2}) \wedge \mathit{Perform}(y, A_{db}, T_{b1}, T_{b2})$

Assume that $\neg R$ is a causal link, $\mathit{cLink}(\neg R, A_{da}, A_{db})$, between A_{da} and A_{db}, negated by an effect or being a precondition R of action A_{dk}. Action A_{dk} from y's point of view cannot be adopted when it is scheduled in the interval of time between the two actions A_{da} and A_{db}. Further, using Arg_p agent x may justify A_{dk} when its effects provide the preconditions for a subsequent action A_{dj}.

Definition 6 (*Arg*$_n$-Norms) *Argument* Arg_n *deals with norm violations caused by adopting or dropping actions. It describes the premises that activate the norm and the conclusions enforced.*

- Given $\mathit{Hold}(y, \mathcal{N}_{Prem}, T_p) \wedge \mathit{Perform}(y, A_{dp}, T_{p1}, T_{p2})$,
- a norm $\mathit{Forbids}$ to $\mathit{Perform}(x, A_{dk}, T_{k1}, T_{k2})/\mathit{Hold}(x, R, T_k)$

$\Rightarrow$ then $\neg\mathit{Perform}(x, A_{dk}, T_{k1}, T_{k2})/\neg\mathit{Hold}(x, R, T_k)$

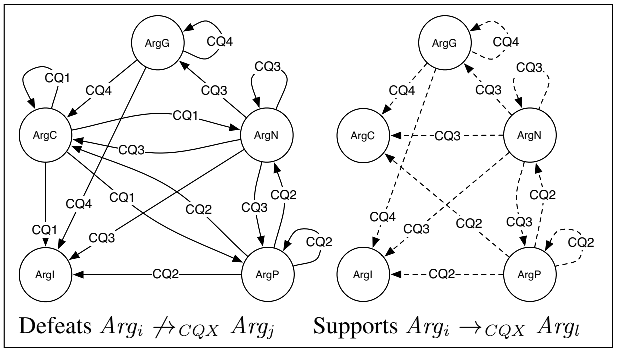
Norms are external regulations about what the agent is obliged to do or forbidden from doing in terms of actions and states of the world. Arg_n is used by y if a proposed action or a state is forbidden. Agent x may also use Arg_n to justify an action or a state if obliged to do that action or achieve that state.

Definition 7 (*Arg*$_g$-Goals) *An argument* Arg_g *is used by a proponent* x *to explore the motivations for adopting goal* ψ_k*.* Arg_g *is also used to describe conflicts between mutually-exclusive goals.*

- Observations $\mathit{Hold}(x, \mathcal{O}_{\psi_k}, T_k)$
- lead to $\mathit{Achieve}(x, \psi_k)$ and $\mathit{Hold}(x, \mathcal{F}_{\psi_k}, T_k)$, [$R_k \in \mathcal{F}_{\psi_k}$]
- lead to $\neg\mathit{Achieve}(y, \psi_j)$ and $\neg\mathit{Hold}(y, \mathcal{F}_{\psi_j}, T_j)$, [$\neg R_k \in \mathcal{F}_{\psi_j}$]

$\Rightarrow$ then $\neg\mathit{Achieve}(x, \psi_k)$ and $\mathit{Achieve}(y, \psi_j)$

In Arg_g the adoption of goal ψ_k is motivated by some observations of the world $\mathcal{O}_{\psi_k} = \{R_a, \ldots, R_l\}$, and ψ_k is achieved when the features of the world $\mathcal{F}_{\psi_k} = \{R_b, \ldots, R_m\}$ hold. Agent y may also argue that its goal ψ_j and x's goal ψ_k are incompatible. For example ψ_j, achieved through $\mathcal{F}_{\psi_j} = \{R_c, \ldots, R_n\}$, negates a feature R_k that must hold to achieve ψ_k ($R_k \in \mathcal{F}_{\psi_k}$ and $\neg R_k \in \mathcal{F}_{\psi_j}$). Here the goals are mutually-exclusive and one agent has to drop its goal.



Defeats $Arg_i \not\rightarrow_{CQX} Arg_j$ Supports $Arg_i \rightarrow_{CQX} Arg_l$

Figure 1. Defeat and Support Relations

Speech Act	Reply: ● Attack ($\mathcal{R}_a$) ⋆ Surrender ($\mathcal{R}_s$)
$propose(Arg_I)$	●$[\mathcal{P}_{contr}, \mathcal{P}_{asym}, \mathcal{P}_{sym}]reject(\ell_{A_{dk}})$ ●$[\mathcal{P}_{asym}, \mathcal{P}_{sym}]why(\ell_{ti})$ and $\ell_{ti} \in Arg_I$ ⋆$[\mathcal{P}_{contr}, \mathcal{P}_{asym}, \mathcal{P}_{sym}]accept(\ell_{A_{dk}})$
$reject(\ell_{A_{dk}})$	●$[\mathcal{P}_{asym}, \mathcal{P}_{sym}]argue(Arg_i)$ and $\{Arg_i \rightarrow_{CQX} Arg_I\} \in \mathcal{R}_{sup}(Arg_I)$ ●$[\mathcal{P}_{asym}, \mathcal{P}_{sym}]why(\neg\ell_{ti})$ and $\ell_{ti} \in Arg_I$ ⋆$[\mathcal{P}_{contr}, \mathcal{P}_{asym}, \mathcal{P}_{sym}]withdraw(\ell_{A_{dk}})$
$withdraw(\ell_{A_{dk}})$	
$accept(\ell_{A_{dk}})$	
$argue(Arg_i)$	●$[\mathcal{P}_{asym}, \mathcal{P}_{sym}]why(\ell_t)$ and $\ell_t \in Arg_i$ ●$[\mathcal{P}_{asym}, \mathcal{P}_{sym}]argue(Arg_j)$ and $\{Arg_j \not\rightarrow_{CQX} Arg_i\} \in \mathcal{R}_{def}(Arg_i)$ ●$[\mathcal{P}_{sym}]argue(Arg_j), \{Arg_j \rightarrow_{CQX} Arg_l\} \in \mathcal{R}_{sup}(Arg_l), Arg_l \in d_i = (m_1, \ldots, m_i)$ ⋆$[\mathcal{P}_{contr}, \mathcal{P}_{asym}, \mathcal{P}_{sym}]withdraw(\ell_{A_{dk}})$ ⋆$[\mathcal{P}_{contr}, \mathcal{P}_{asym}, \mathcal{P}_{sym}]accept(\ell_{A_{dk}})$
$why(\ell_{ti})$	●$[\mathcal{P}_{asym}, \mathcal{P}_{sym}]argue(Arg_j)$ and $\ell_{ti} \in Arg_j$ ⋆$[\mathcal{P}_{contr}, \mathcal{P}_{asym}, \mathcal{P}_{sym}]withdraw(\ell_{A_{dk}})$ ⋆$[\mathcal{P}_{contr}, \mathcal{P}_{asym}, \mathcal{P}_{sym}]accept(\ell_{A_{dk}})$

Table 1. Speech acts for each protocol

Defeat and support relations. Here we propose a formalisation of relations among arguments, focussing on what arguments defeat or support others. The argumentation schemes presented include critical questions that allow agents to select new arguments according to their commitments and how they conflict with actions, goals and norms. The critical questions guide an agent to internally identify counter-arguments that defeat the opponent's view or arguments that support its position. Here we consider the following critical questions, which are pointers to the arguments expressed in brackets: "**CQ1:** Is the action possible given other concurrent actions in the plan?" (Arg_c); "**CQ2:** Is the action possible according to causal plan constraints?" (Arg_p); " **CQ3:** Is there any norm that regulates actions or states of the world?" (Arg_n); " **CQ4:** Does the goal conflict with another goal?" (Arg_g). The structure of the arguments can be seen as a set of premises followed by conclusions ("⇒then").

Definition 8 (Defeat relations) *A defeat relation is represented as* $Arg_i \not\rightarrow_{CQX} Arg_j$ *where* $Arg_i \in Args$ *defeats* $Arg_j \in Args$ *identified by the critical question* CQX*. In each relation the argument* Arg_i *has conclusions that negate a premise or a conclusion of argument* Arg_j*. The set of defeat relations is defined as* $\mathcal{R}_{def} = \{(Arg_j, Arg_i) : \forall Arg_i, Arg_j \in Args \text{ where } Arg_i \not\rightarrow_{CQX} Arg_j\}$

When an argument Arg_j is presented by an agent, Arg_i is a move from the opponent who criticises an action, a goal or a feature of the world. $\mathcal{R}_{def}(Arg_j) \subseteq \mathcal{R}_{def}$ identifies the defeat relations of Arg_j.

Definition 9 (Support relations) *A support relation is represented as* $Arg_i \rightarrow_{CQX} Arg_l$ *where* $Arg_i \in Args$ *justifies a premise or a conclusion of the argument* $Arg_l \in Args$ *identified by a critical question* CQX*. The set of support relations is defined as* $\mathcal{R}_{sup} = \{(Arg_j, Arg_i) : \forall Arg_i, Arg_j \in Args \text{ where } Arg_i \rightarrow_{CQX} Arg_j\}$

Arguments Arg_i and Arg_l are different moves of the same agent attempting to defend an action, a goal or a feature of the world in its plan. Valid defeat and support relations are represented in Figure 1.

2.2 Dialogue System

The dialogue system used is an instantiation of that proposed by Kok et al. [3]. Here we describe only the characteristics of the dialogue pertaining to our work. A dialogue system is a tuple $\mathcal{D} = \langle \mathcal{L}_c, \mathcal{P}, \mathcal{C} \rangle$ where $\mathcal{L}_c$ is the communication language, $\mathcal{P}$ the communication protocol and $\mathcal{C}$ a set of rules specifying the effects of the locutions. A dialogue composed by h moves is represented as a set $d_h = (m_1, \ldots, m_h)$. For each move m_i the identifier, the player, the speech act and the target of the move are defined. The move m_h refers to the last move of dialogue d_h. The communication language is a tuple $\mathcal{L}_c = \langle \mathcal{L}_p, \mathcal{R}_a, \mathcal{R}_s \rangle$ where $\mathcal{L}_p$ is a set of speech acts, $\mathcal{R}_a$ (attacks) and $\mathcal{R}_s$ (surrenders) are binary relations in $\mathcal{L}_p \times \mathcal{L}_p$. Each speech act in $\mathcal{L}_p$ is the form $p(\vartheta_i)$ where p is a performative and ϑ is a sentence in the topic language $\ell_{ti} \in \mathcal{L}_t$ or an argument $Arg_i \in Args$. In particular, we refer to the proposal as a sentence $\ell_{A_{dk}} = Perform(x, A_{dk}, T_{k1}, T_{k2})$.

The speech acts used here are: $propose(Arg_I)$ for proposing an action ($\ell_{A_{dk}} \in Arg_I$); $accept(\ell_{A_{dk}})$ and $reject(\ell_{A_{dk}})$ for accepting or rejecting the proposal; $withdraw(\ell_{A_{dk}})$ for withdrawing the proposal; $why(\ell_{ti})$ is a question "why perform/achieve?" to challenge the performance of an action or the achievement of a state; $argue(Arg_i)$ for presenting an argument.

The dialogue system is designed for two agents. Initially each agent creates an individual plan. The proponent starts the dialogue proposing an action from its plan to the other agent, and the dialogue progresses in a turn-taking fashion. When an agent passes, the proponent withdraws its proposal or the opponent accepts it, the dialogue terminates. On termination, agents may re-plan taking into account new information acquired during the dialogue. Table 1 summarises the speech acts and possible responses according to the protocol.

For evaluation purposes, two assumptions are made: if an agent concedes it must concede the original proposal and the dialogue terminates; and agents may only discuss a single issue at a time. When the dialogue has terminated and the agents have re-planned they start a new discussion about new commitments if necessary. Thus, the protocol does not include the possibility for agents to offer an alternative. This allows us to map the information shared to the solution identified, and hence to better assess the re-planning process.

In this research we claim that the information shared during the dialogue leads agents to identify better collaborative plans. To test this claim, we consider three dialogue protocols $\mathcal{P}_{contr}$, $\mathcal{P}_{asym}$, and $\mathcal{P}_{sym}$ that correspond to different degrees of freedom in moving arguments. Protocol $\mathcal{P}_{contr}$ is a control condition where agents are not permitted to exchange arguments other than accepting or rejecting the claim. The protocols that use the argumentation schemes discussed are a symmetric protocol ($\mathcal{P}_{sym}$) where the proponent and opponent may

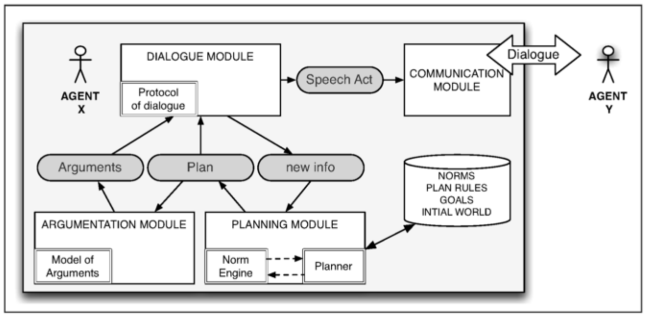


Figure 2. System Architecture.

use defeat or support relations to form arguments, and an asymmetric protocol ($\mathcal{P}_{asym}$) where the opponent explores its objections to the proposed action which are defended by the proponent. The latter is analogous to rigorous persuasion (RPD) and the former akin to permissive persuasion (PPD) as discussed by Walton and Krabbe [8]. In protocol $\mathcal{P}_{asym}$ an opponent can only attack the last move m_h of the proponent, or it surrenders. This induces asymmetry since often an opponent attacks the proponent without sharing enough information about plan conflicts. Hence, the proponent does not have the means to challenge the opponent's position and it may only defend its claim. In protocol $\mathcal{P}_{sym}$, agents must respect their turns but they may move an argument in support of another previously moved. In a dialogue $d_h = (m_1, \ldots, m_h)$ if the next turn m_{h+1} is of agent y, agent y may move $m_{h+1} = argue(Arg_l)$. Argument Arg_l supports an argument Arg_i, the subject of the move m_i where $1 \leq i < h$. The resulting dialogue is symmetric in the sense that both agents may attack others' arguments or they may, at any time, take the role of a proponent and defend an argument they introduced in the dialogue.

3 System Architecture

The agents' planning domain concerns operations of a local authority, x, and a humanitarian organisation, y, for evacuation of people after a disaster. Initially each agent creates an individual norm-consistent plan. When an individual plan is specified, an agent checks if there are actions in its plan that require discussion. The dialogue module selects the next legal move according to the protocol and the arguments available are generated by the argumentation module. However, agents do not follow a preferential order for exchanging arguments. Figure 2 shows an overall view of the system.

The identification of plan alternatives is made through a three-steps dynamic process that considers conflicting intervals of actions identified during the dialogue. To illustrate the process, consider a situation where agents discuss the repair of the water supply in location $locA$. Agent x proposes to stop the water supply in $locA$ from T_2 to T_3. Agent y argues about the need for water in that location for performing other activities that include building and running a field hospital respectively between $[T_3, T_5]$ and $[T_5, T_9]$. However, agent x has a norm that obliges it to stop the supply because the water is contaminated and, thus, it is dangerous for people. Further, $locA$ is not safe since y has scheduled the use of excavators at that location from T_2 to T_{20}, thus y is forbidden from running the hospital. The discussion ends with y conceding and re-elaborating the conflicting part of its plan. From the dialogue, y establishes the conflicting intervals of building the field hospital and running it resulting $[T_2, T_3]$ and $[T_2, T_{20}]$ respectively. The first attempt of re-planning consists of moving both actions before T_2. Given that the planning period begins at T_0, this option is not possible. The second option is to build the field hospital after T_3 and start using it after T_{20}. However, assuming that y has an internal constraint whereby the hospital has to run at T_{15}, this is not a good solution. Thus, agent y may decide to build the field hospital in a different location $locB$ at time T_2. This would fulfil its requirements as well as x's requirements and norms. Assuming that no solutions were available, agent y may have been forced to drop the goal that require building and running a field hospital and all the actions involved to achieve them. In the control condition, $\mathcal{P}_{contr}$, the argumentation module is disabled and agents attempt re-planning with random intervals.

4 Evaluation

In this research, the metric for evaluation is the feasibility of the resulting collaborative plans. Feasibility is measured by the number of conflicts of different types between individual plans that can hamper the execution of interdependent tasks. Here we detail the experimental design used in this empirical study and analyse our results.

Experimental Design. Individual plans assigned to agents have an average of 5 objectives and 50 actions, and are internally consistent with respect to the agent's beliefs about the initial state. Agents are not given information about conflicts between their plans. Collaboration is enforced for a total of 5%-10% of actions in each plan. We randomly generated 40 different initial plans for each agent with variations in time-scale, norms, goals and beliefs about the initial state. In order to represent sufficient conflicts between pairs of plans, we consider 11 conflict configurations in a range from 25 to 75 initial conflicts (i.e. 25,30,…,70,75). Among the 40 × 40 initial plans, 25 pairs where selected for each different conflict configuration; a total of 275 pairs of plans. Each pair was used in an run for the three protocols: $\mathcal{P}_{contr}$, $\mathcal{P}_{sym}$, $\mathcal{P}_{asym}$. The plans for each experimental run were analysed to measure the conflicts before and after the dialogue.

These sets are used for the analysis of results: $\boldsymbol{\mathcal{C}_{tot}}$ conflicts before the experiment (complexity of the problem); $\boldsymbol{\mathcal{C}_{end}}$ conflicts after the experiment; $\boldsymbol{\mathcal{C}_{add}}$ new conflicts added during the dialogue such that $\mathcal{C}_{start} \cap \mathcal{C}_{add} = \emptyset$; $\boldsymbol{\mathcal{C}_{rem}}$ conflicts removed during the dialogue; $\mathcal{C}_{rem} = \mathcal{C}_{start} \setminus (\mathcal{C}_{end} \setminus \mathcal{C}_{add})$; $\boldsymbol{\mathcal{C}_{drop}}$ conflicts removed by dropping goals; $\boldsymbol{\mathcal{C}_{solv}}$ conflicts solved where $\mathcal{C}_{sol} = \mathcal{C}_{rem} \setminus \mathcal{C}_{drop}$. The set of added conflicts $\mathcal{C}_{add}$ includes obligations not fulfilled because of removed actions. The conflicts in C_{drop} are treated as unsolved because they are only removed due to a goal being dropped.

Results. We claim in this research that our model of deliberation dialogue based on argumentation schemes for the identification of plan conflicts will lead agents to identify more feasible collaborative plans using a dynamic re-planning mechanism which takes into account new information acquired during the dialogue. Thus, the hypotheses for evaluation are the following: *(i)* the use of argumentation schemes for structuring deliberative dialogue increases the number of conflicts resolved between interdependent plans; *(ii)* the use of argumentation schemes reduces the need for agents to drop goals for resolving conflicts; *(iii)* the use of a symmetric protocol is more effective in conveying information about conflicts between plans.

In the figures, lines represent linear regression models and points represent the data for the identification of these trends.

Figure 3.A reports the conflicts removed with different protocols ($|\mathcal{C}_{rem}| \sim |\mathcal{C}_{tot}|$). The graph shows that the protocols have a similar performance in this regard. We performed a statistical significance test of conflicts removed against the total number of conflicts for the three different protocols. The test revealed that there is insufficient

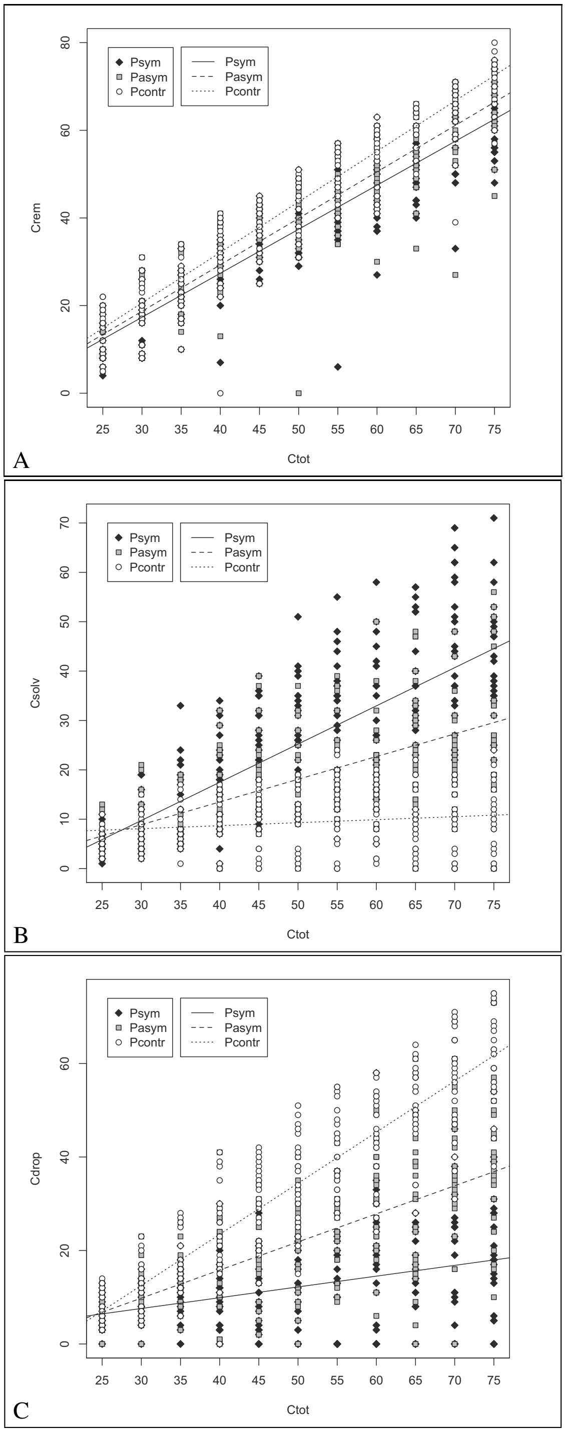


Figure 3. Conflicts removed, conflicts solved, and conflicts dropped as total conflicts increases.

evidence to conclude that the conditions are different ($p = 0.112$).

In Figure 3.B we plot the number of conflicts solved (i.e. removed but not due to a goal being dropped) as the complexity of the problem increases ($|\mathcal{C}_{solv}| \sim |\mathcal{C}_{tot}|$). The graph shows that the conflicts solved are significantly higher when argumentation schemes are used to guide the dialogue ($\mathcal{P}_{sym}$ and $\mathcal{P}_{asym}$) than in the control condition ($\mathcal{P}_{contr}$). In $\mathcal{P}_{contr}$ the trend is almost flat, demonstrating that agents solve a similar number of conflicts regardless of the complexity of the problem. This result provides evidence for the hypothesis that many conflicts may only be solved through the exchange of arguments. Using argumentation-based protocols agents are able to share relevant information about existing plan, norm and goal commitments.

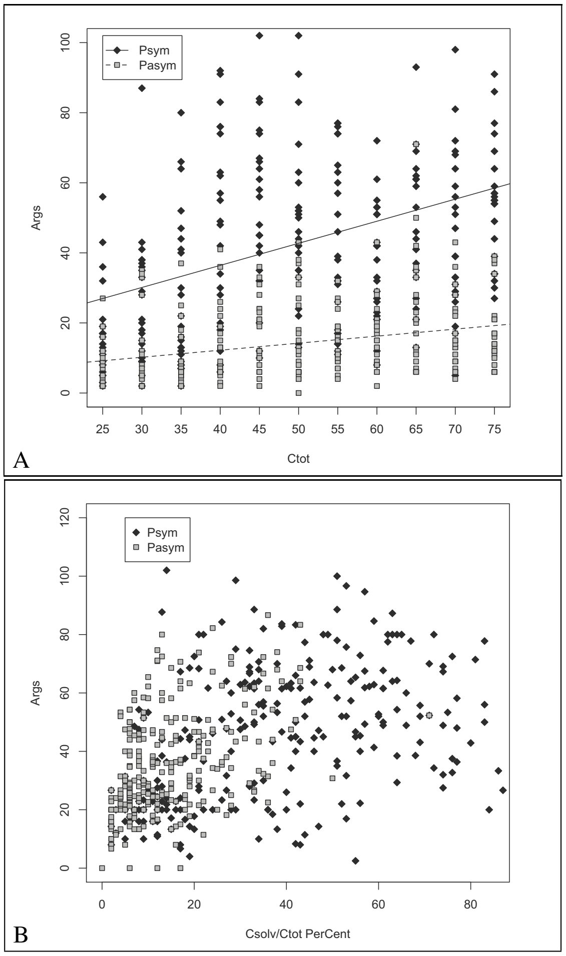


Figure 4. Analysis of Arguments

Figure 3.C presents the number of conflicts removed because of goals being dropped for each run of the three protocols ($|\mathcal{C}_{drop}| \sim |\mathcal{C}_{tot}|$). The trends show that protocol $\mathcal{P}_{contr}$ has a higher rate of conflicts resolved due to goals being dropped compared to the argumentation-based protocols. This result shows that agents employing $\mathcal{P}_{contr}$ are not able to find effective solutions for the conflicts discovered and in order to resolve conflicts, they have to drop individual goals. In $\mathcal{P}_{asym}$ and particularly in $\mathcal{P}_{sym}$ the trend is significantly lower, supporting the hypothesis that agents using argumentation schemes are able to identify feasible solutions for conflicts among interdependent plans avoiding the need to drop goals.

The last hypothesis refers to the difference in the performance of the asymmetric $\mathcal{P}_{asym}$ and the symmetric $\mathcal{P}_{sym}$ protocols. Although the two argumentation protocols present a similar trend for conflicts solved and dropped, $\mathcal{P}_{sym}$ performs consistently and significantly better than $\mathcal{P}_{asym}$. Figure 5 presents the proportion of arguments of different types used on average in each of the protocols. Figure 4.A shows that the number of arguments exchanged increases as the complexity of the plan increases ($|Args| \sim |\mathcal{C}_{tot}|$). Further, the total number of arguments exchanged in $\mathcal{P}_{sym}$ tends to be significantly

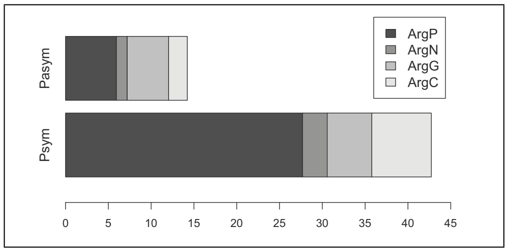


Figure 5. Proportion of argument types used on average per run.

higher than with $\mathcal{P}_{asym}$. In $\mathcal{P}_{asym}$ the proponent only defends attacks on its claim as the opponent criticises it. Dialogues with $\mathcal{P}_{sym}$ tend to be longer than $\mathcal{P}_{asym}$. This is because in $\mathcal{P}_{sym}$ both agents play the opponent and proponent role permitting additional information to be exchanged; e.g. justifications for an agent's commitment. Thus, $\mathcal{P}_{sym}$ enables agents to explore the causes of the conflicts between the plans in more depth. Figure 4.B presents the relationship between the conflicts solved and the number of arguments exchanged ($|Args| \sim |\mathcal{C}_{solv}|/|\mathcal{C}_{tot}|$). Points towards the right hand side of the graph indicate higher ratios of conflicts solved. This is principally populated by problems discussed with $\mathcal{P}_{sym}$. The graph shows that the ratio of conflicts solved is higher in $\mathcal{P}_{sym}$ and this has been achieved by exchanging a higher number of arguments. This result provides evidence for there being a tradeoff in practice between the complexity of the dialogue and the number of conflicts that can be solved. We conclude that, although $\mathcal{P}_{sym}$ leads to more complex dialogues, it is more effective in conveying information about conflicts and in resolving complex interdependencies between agents' plans.

In these experiments we studied the impact of increasing the total number of conflicts in the three protocol conditions. The effects on the conflicts solved, conflicts dropped and arguments exchanged were found to be statistically significant using two-way ANOVA tests at $p < 0.05$. Using post-hoc analysis, the results for the three protocols were significantly different from each other.

5 Discussion and Conclusion

In this research we have considered complex planning problems where agents have to fulfil individual objectives, respecting plan constraints as well as normative ones. However, when agents collaborate to accomplish certain tasks, they must form an agreement on a shared plan [4]. This leads agents to engage in dialogue focussed on resolving conflicting opinions on what to do. In recent years several alternative models of argumentation for deliberative dialogues have been proposed [1, 6, 7] providing additional heuristics for structuring deliberative dialogue. In this paper we go further by empirically evaluating the model presented in [7]. We have demonstrated that agents employing appropriate argumentation-schemes are indeed able to agree more effective collaborative plans.

In previous research, only few approaches have conducted an evaluation of the benefits of argumentation schemes. Karunatillake et al. [2] demonstrated that the use of argumentation schemes enables agents to identify a better allocation of limited resources within a society. However, this work considers relatively simple domains, while we address complex deliberation problems to resolve conflicts between interdependent plans. To date, the potential of argumentation schemes in deliberative dialogue has been investigated exclusively through the use of examples [1, 6], where agents choose their solutions among a predefined set of alternatives. Here we have presented a system where agents construct alternatives dynamically according to the information acquired during the dialogue. This enabled an evaluation of how information shared through the argumentation schemes affects the re-planning. The results show that the use of appropriate argumentation schemes provides an effective method to share relevant information for solving plan and normative conflicts.

We considered here agents that choose actions with the aim of reducing the number of conflicts between interdependent plans. However, existing argumentation-based models have proposed methods for comparing arguments; i.e. actions chosen based on their utility in the plan [6], costs and social implications in acquiring resources [2], or social values of the partners involved [1]. In future work we will consider the selection of arguments to exchange during the dialogue according to their importance; i.e. agents may perform a quantitative analysis of costs and benefits of actions, and penalties for norm violation. We aim to assess whether the information shared during the dialogue enables agents to effectively solve more critical conflicts.

In conclusion, in this paper we outlined an argumentation-based model of deliberative dialogue grounded upon argumentation schemes focussed on identifying goal, norm and plan conflicts. We presented a method for assessing how the information shared during the dialogue affects the conflict resolution process. Our study showed that the use of argumentation schemes enhances the ability of agents to resolve conflicts between interdependent plans avoiding the need to drop individual goals. We have demonstrated that a symmetric protocol, albeit resulting in more complex dialogues, is more effective in conveying information about conflicts between plans and this information enables the identification of more feasible collaborative solutions. Thus, the use of argumentation schemes in deliberative dialogue is an effective mechanism to establish agreements on how to act together.

ACKNOWLEDGEMENTS

This research is supported by the award made by the RCUK Digital Economy programme to the dot.rural Digital Economy Hub; award reference: EP/G066051/1.

REFERENCES

[1] K. Atkinson and T. Bench-Capon, 'Practical reasoning as presumptive argumentation using action based alternating transition systems', *Artificial Intelligence*, **171**(10-15), 855–874, (2007).

[2] N. C. Karunatillake, N. R. Jennings, I. Rahwan, and P. McBurney, 'Dialogue games that agents play within a society', *Artificial intelligence*, **173**(9-10), 935–981, (2009).

[3] E. M. Kok, J. J. C. Meyer, H. Prakken, and G. A. W. Vreeswijk, 'A formal argumentation framework for deliberation dialogues', in *Proceedings of the 7th International Workshop on Argumentation in Multi-Agent Systems*, (2010).

[4] S. Kraus, K. Sycara, and A. Evenchik, 'Reaching agreements through argumentation: a logical model and implementation', *Artificial Intelligence*, **104**(1-2), 1–69, (1998).

[5] J. A. Pinto and R. Reiter, 'Reasoning about time in the situation calculus', *Annals of Mathematics and Artificial Intelligence*, **14**, 251–268, (1995).

[6] I. Rahwan and L. Amgoud, 'An argumentation-based approach for practical reasoning', in *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 347–354, (2007).

[7] A. Toniolo, T. J. Norman, and K. Sycara, 'Argumentation schemes for collaborative planning.', in *Proceedings of the 14th International Conference on Principles and Practice of Multi-Agent Systems*, (2011).

[8] D.N. Walton and E.C.W. Krabbe, *Commitment in dialogue: Basic concepts of interpersonal reasoning*, State University of New York Press, 1995.