Argumentation-based Negotiation in t-DeLP-POP[1]

Pere PARDO [a,2], Pilar DELLUNDE [a,b] and Lluís GODO [a]

[a] *IIIA-CSIC, Institut d'Investigació en Intel·ligència Artificial*

[b] *Universitat Autònoma de Barcelona*

Abstract. In this contribution, we propose to model argumentation-based negotiation in terms of t-DeLP-POP, a partial order planning system that incorporates temporal defeasible logic. This logic combines temporal facts and durative rules into temporal arguments. We propose a dialogue protocol for the negotiation of plans in this planning system that models a variety of scenarios for argumentative negotiation of complex services. Then we consider case studies in the literature can be naturally modeled by dialogues in this logic-based planning framework.

Keywords. Negotiation, Argumentation, Planning, Temporal Defeasible Logic.

1. Introduction

Negotiation skills are an important ability for autonomous agents in decentralized multi-agent systems. Agents pursuing their own goals are inter-dependent, specially if abilities, rights or resources are unequally distributed among them.

Traditionally, purely quantitative approaches to negotiation are somewhat limited in expressivity, since agents merely exchange offers and accept/reject messages, instead of communicating what is good or bad with an offer, and why. Argumentation-based negotiation (ABN) is a recent area of research that tries to overcome these limitations along this line and speed up the process of reaching agreements.

In this contribution we study a logic-based planning framework as a foundation for ABN, with a focus, as in [14], on the agent architecture. The present framework aims at a two-fold integration of argumentation within the mental model of the negotiating agent: descriptive argumentation is based on the *logical program* of an agent (i.e. on her beliefs: facts and rules; while practical argumentation builds upon the *planning domain* of some agent (beliefs, actions and goals). An agent's plan search also involves taking part in several concurrent, pairwise ABN dialogues with other agents. Communicated content, though, may propagate to dialogues involving other agents. In any case, communications are determined by the agents' beliefs and concession functions.

[1]The authors acknowledge partial support of the Spanish MICINN project CONSOLIDER-INGENIO 2010 Agreement Technologies CSD2007-00022; LoMoReVI FFI2008-03126-E/FILO (FP006); ARINF TIN2009-14704-C03-03 and the Generalitat de Catalunya grant 2009-SGR-1434.

[2]Corresponding Author: P. Pardo, IIIA-CSIC, Campus UAB s/n, Bellaterra, Catalonia (Spain); E-mail: pardo@iiia.csic.es

A negotiation dialogue starts once a customer communicates some of her goals (say, *be at Oslo at* $t+5$). The t-DeLP-POP planning system models *negotiable services* as actions that can be used to solve these goals by refining the customer's plans with them (e.g. *buy train tickets to Oslo at* t). Thus, planning actions are the atomic objects of negotiation and offers consist of a pair of such actions to be exchanged. Arguments are also ways to enforce an open goal (i.e. as the argument's conclusion), just like actions (as one of its effects), and as such, they can form as an offer t-DeLP arguments represent causal or temporal processes, with premises and conclusion being, resp., causes and effect. These processes may become triggered -intentionally or not- by the execution of planned actions (plus initial facts). Thus, within a negotiation it can be argued whether the plan-triggered processes will actually succeed w.r.t. the goals (e.g. the expected process from *being at the train to Oslo at* $t+1$ to *be at Oslo at* $t+5$ can be threatened by a *snow storm occurred at* t).

Speech acts of ABN dialogues between negotiating agents consist of:

- proposals (demands; offers or acceptance), encoded as goals, resp., as plan steps,
- theoretical arguments (plan threats, agent threats), encoded as argument threats,
- practical arguments (persuasion; challenge), as actions; resp., planning domains.

The present t-DeLP-POP-based model of ABN makes use of the idealized assumption: agents are totally honest and trust each other.[3]

The paper is structured as follows: we briefly introduce t-DeLP-POP in the Preliminaries, and discuss basic issues in its multi-agent extension. Then we present the negotiation framework and the protocol for ABN dialogues. Finally, we model and discuss in our framework some examples from the ABN literature.

Related Work. Our approach is a temporal extension of defeasible logic DeLP [6], and its combination with partial order planning (POP) in [7]. A multi-agent extension of the latter was studied in [10] for the cooperative case. Here we explore a more general case: including non-cooperative scenarios too, and based on the temporal defeasible logic t-DeLP [11], and its combination with POP [9].

In the literature, some proposals for ABN exist based on general argumentation frameworks, of argumentation for ABN [2,3,5], (that also include agent threats). For ABN with discussion of goals and services, proposals based on some modal (epistemic, dynamic) logic exist, like [16], [8] or the (multi-context) BDI logic [14]; see also [15], [4] for negotiation protocols). While the modal logics can express nesting of mental attitudes (beliefs, etc.,) they also present limitations due to monotonicity (a less convenient representation of actions).

2. Preliminaries: t-DeLP-POP

t-DeLP is a temporal extension of the DeLP argumentation framework proposed in [11]. We take $\mathbb{N}$ as our working set of discrete time-points. The language consists of temporal literals and rules. Temporal literals are of the form $\ell = \langle p, t \rangle$ or $\ell = \langle \sim p, t \rangle$ and express "p (resp. not-p) holds at time t", from a given set of variables $p \in \mathsf{Var}$. Strong negation,

[3]Thus, relevant information, i.e. arguments, are freely given regardless of the consequences w.r.t. one's interests. The necessary additions or modifications of the present approach lie out of the scope of this paper.

denoted $\sim p$, extends to literals: $\sim\ell = p$ if $\ell = \sim p$ and $\sim\ell = \sim p$ if $\ell = p$; and also to sets: $\overline{X} = \{\sim\ell \mid \ell \in X\}$. Defeasible rules δ are of the form $\langle p, t\rangle \prec \langle p_0, t_0\rangle, \ldots, \langle p_n, t_n\rangle$, satisfying $t \geq \max\{t_0, \ldots, t_n\}$. Such rules read: premises $\langle p_0, t_0\rangle, \ldots$ constitute in principle a cause for (or a reason for) $\langle p, t\rangle$; thus, $\mathsf{body}(\delta) := \{\langle p_0, t_0\rangle, \ldots\}$ precede or occur no later than its conclusion $\mathsf{head}(\delta) := \langle p, t\rangle$. A special type of rules are *persistence* rules δ_p of the form $\langle p, t+1\rangle \prec \langle p, t\rangle$, stating p in principle will be preserved from t to $t+1$. Knowledge of p *holds at* t is represented as a strict rule with empty body $\langle p, t\rangle \leftarrow$ in the set of initially known facts Ψ, but will be denoted simply as $\langle p, t\rangle$.

A temporal defeasible logic program (t-de.l.p.) is a pair (Ψ, Δ) where Ψ is a consistent set of basic facts (i.e. no pair $\langle p, t\rangle$, $\langle \sim p, t\rangle$ is in Ψ) and Δ is a set of temporal defeasible rules. The notion of *derivability of literals* in a t-de.l.p. (Ψ, Δ) is defined by closure of Ψ under *modus ponens* with Δ rules.

Derivability is monotonic and, typically, the literals derivable in a program will form form an inconsistent set, i.e. with some p and $\sim p$ being derivable.[4] The refined, consistent notion of (non-monotonic) logical consequence (called *warrant*) is defined by means of an argumentative process.

An *argument* for a temporal literal $\langle p, t\rangle$ is a $\subseteq$-minimal set of rules $\mathcal{A} \subseteq \Delta$ such that $\mathcal{A} \cup \Psi$ is consistent and that $\langle p, t\rangle$ is derivable from $\mathcal{A} \cup \Psi$ -the latter also denoted $\mathsf{concl}(\mathcal{A}) = \langle p, t\rangle$; while its set of premises is defined by $\mathsf{base}(\mathcal{A}) := \mathsf{body}[\mathcal{A}] \smallsetminus \mathsf{head}[\mathcal{A}]$ and the argument's duration $\|\mathcal{A}\|$ is the difference $t - t'$, where t' is given by the earliest of $\mathcal{A}$'s premises $\langle p', t'\rangle \in \mathsf{base}(\mathcal{A})$. An argument $\mathcal{B}$ for another literal $\langle q, t'\rangle$ is a subargument of $\mathcal{A}$ whenever $\mathcal{B} \subseteq \mathcal{A}$ and $\mathsf{base}(\mathcal{B}) \subseteq \mathsf{base}(\mathcal{A})$; this $\mathcal{B}$ is denoted $\mathcal{A}(\langle q, t'\rangle)$.

An argument $\mathcal{A}$ for $\langle p, t\rangle$ *attacks* another argument $\mathcal{A}'$ when $\mathcal{A}'$ has a subargument $\mathcal{B}$ for $\langle \sim p, t\rangle$. Attacks merely point out the existence of a logical conflict. A preference relation between arguments is therefore needed to decide which arguments prevail. In this temporal framework, we opt for a purely formal criterion of preference for *better information*, based on more information in the premises, or based on information which is more recent (i.e. temporally closer to the time t of the attacked $\mathsf{concl}(\mathcal{B}) = \langle \sim p, t\rangle$).

Finally, a literal $\langle \ell, t\rangle$ is *warranted* in (Ψ, Δ), denoted $\ell \in \mathsf{warr}(\Psi, \Delta)$, if an argument $\mathcal{A}$ for $\langle \ell, t\rangle$ exists and is undefeated by the other arguments. To see whether this is the case, first note defeaters can have defeaters (see Figure 1 (Top Left)); the relation *is defeated by* determines a tree with root $\mathcal{A}$. A marking procedure from leaf nodes (unattacked arguments are undefeated) up to $\mathcal{A}$ is defined by the condition: $\mathcal{B}$ is defeated iff it is defeated by an undefeated argument $\mathcal{C}$ in the tree (see Figure 1 (Bottom Left)).

A basic property of t-DeLP is that it prunes inconsistencies out of any t-de.l.p.:

Theorem 1 *[11] For any t-de.l.p.* (Ψ, Δ)*,* $\mathsf{warr}(\Psi, \Delta)$ *is consistent.*

t-DeLP-POP [9] is a planning system that combines backward search in partial order planning (POP) with forward reasoning by means of t-DeLP to compute plan progression. A feature of this system is that the representation of a an action splits into a deterministic action and an associated "non-monotonic theory", encoded within Δ, capturing effects that are context-dependent[5].

[4] Following the example in section 1, the basic facts (*be at the train to Oslo at* $t+1$ and *a snow storm occurred at* t) suffice to trigger a rule concluding that *[I will] be at Oslo at* $t+5$ and a rule for its negation.

[5] Thus we split a classical planning action *go by train to Oslo* into: (1) an action $\mathsf{get.on.train.Oslo(t)}$ with effect $\langle\mathsf{got.on.train.to.Oslo, t}\rangle$; (2) a rule from this effect to $\langle\mathsf{be.at.Oslo, t+4}\rangle$, and an exception rule for its negation further based on the the fact $\langle\mathsf{snow.storm.at.Oslo, t}\rangle$.

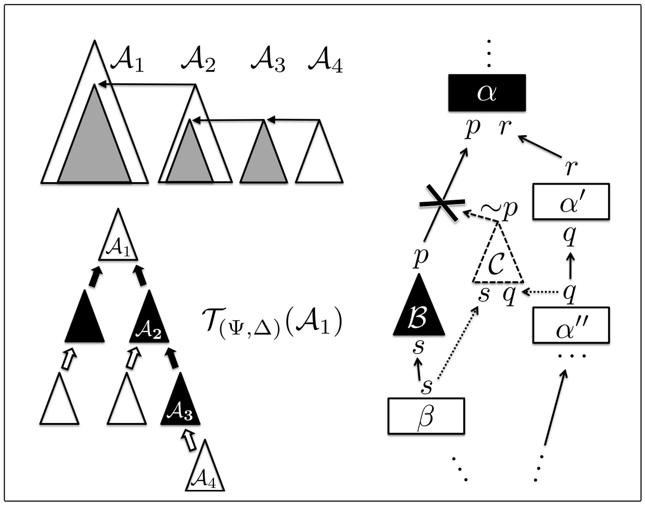

Figure 1. (Top Left) A line of defeaters, with attacked sub-arguments in grey. (Bottom Left) The dialectical tree for $\mathcal{A}_1$, with defeated arguments in black. (Right) A potential (uninteded) argument threat $\mathcal{C}$ for $\mathcal{B}, \alpha$.

A *planning domain* is a tuple $((\Psi, \Delta), A, G)$ for beliefs, actions and goals, representing an agent who wants to find a solution plan: a sequence of actions leading the current Ψ-state into a state where goals G are achieved. As in POP, plans are incrementally built as a set of refinement steps until a (optimal) solution is found. During plan search the planner agent does only impose the minimal constraints on the execution ordering of planned actions, which form a partial order. Actions $\alpha = (\mathsf{P}(\alpha), \mathsf{X}(\alpha), \mathsf{cost}(\alpha))$ have a duration denoted $\|\alpha\| \in \mathbb{N}$, and consist of preconditions $\mathsf{P}(\alpha)$ holding at time $\mathbf{t}(\alpha)$ (a variable), and effects $\mathsf{X}(\alpha)$ holding at $\mathbf{t}(\alpha) + \|\alpha\|$; among these, we assume by default a dummy effect α'ed $\in \mathsf{X}(\alpha)$ stating α *was executed*. The cost $\mathsf{cost}(\alpha)$ is some positive real number. Action α is also denoted $\mathsf{X}(\alpha) \overset{\alpha}{\longleftarrow} \mathsf{P}(\alpha)$. Ψ and G are encoded as dummy actions $\alpha_{\Psi_t} = (\emptyset, \{\langle \ell, t\rangle \in \Psi\}, 0)$ (for each t) and $\alpha_G = (G, \emptyset, 0)$ with no duration or cost; and related by the constraint: α_{Ψ_0} *occurs before* α_G.

Rules Δ are temporal abstractions of their t-DeLP counterparts: we have now general rules of the form $\delta = \ell \prec (\ell_0, d_0), (\ell_1, d_1), \ldots, (\ell_n, d_n)$, where now (ℓ_n, d_n) denotes the delay (i.e. rule $\langle p, 4\rangle \prec \langle q, 1\rangle, \langle r, 2\rangle$ is an instance of general rule $p \prec (q, 3), (r, 2)$). General rules δ combine into argument steps $\mathcal{A} \subseteq \Delta$, using variables $\langle p_i, \mathbf{t}(\mathcal{A}) + n\rangle$ instead of d_i-values in (p_i, d_i). The duration $\|\mathcal{A}\|$ of $\mathcal{A}$ is defined by the maximum of sums of d_j's in a path of rules from $\mathsf{base}(\mathcal{A})$ to $\mathsf{concl}(\mathcal{A})$. Then n is just $\|\mathcal{A}\|$ minus the sum of the d_js (incl. d_i) in the path of rules from p_i to $\mathsf{concl}(\mathcal{A})$.

A threat can be an argument step, or an argument that is triggered as a side-effect of the plan -as in Figure 1 (Right). The former can only be resolved by reordering the threat $\mathcal{B}$ to the future; for the latter one can also impose arguments defeating the threat.

A plan is a triple $\Pi = (A_\Pi, \mathsf{Goals}(\Pi), \mathbb{I}(\Pi))$ with actions used (non-concurrently), pending goals and inequations expressing constraints on the temporal variables $\mathbf{t}(\kappa)$. The cost of a plan is $\Sigma_{\alpha \in A_\Pi} \mathsf{cost}(\alpha)$. For dummy actions, we may have $\{0 \leq \mathbf{t}(\alpha_G), \mathbf{t}(\alpha_G) \leq 100\} \subseteq \mathbb{I}(\Pi)$, the latter imposing a deadline of 100 time units for any solution plan. Solving a goal in $\mathsf{Goals}(\Pi)$ consists in adding a new constraint to $\mathbb{I}(\Pi)$: imposing the plan step κ before the step κ' whose $\mathsf{base}(\cdot)$ or $\mathsf{P}(\cdot)$ set contain the literal being solved:

add $\mathbf{t}(\kappa) + \|\kappa\| \leq \mathbf{t}(\kappa')$. Moving a threat to the future, κ' being the step supported by the threatened step). A non-deterministic search algorithm for the space of plans consists in:

0. Start with the empty plan $\Pi_\emptyset = (\emptyset, G, \mathbb{I}(\Pi_\emptyset))$.
1. If an unsolved goal or threat exists, choose a goal or threat. Otherwise terminate.
2. Choose some action- or argument-step, or resp. some threat resolution move (if Π cannot be refined, then backtrack to the parent node). Refine Π.
3. Update the set of unsolved goals and threats. Go to step 1.

Given a solution plan Π for $((\Psi, \Delta), A, G)$, any sequential execution of A_Π, given by a model $\tau : A_\Pi \to \mathbb{N}$ of the inequations $\mathbb{I}(\Pi)$ will enforce G according to t-de.l.p. defined by Ψ-plus-actions' effects $\mathsf{X}(\tau(\alpha))$, for any $\alpha \in \Pi$ (and conditional on their preconditions being warranted). The results are stronger if $\mathbf{t}(\alpha_G)$ is imposed a bound.

Theorem 2 *[9] The search algorithm is correct; under temporal deadlines $\mathbf{t}(\alpha_G) \leq k$, the algorithm is complete. Uniform cost heuristic for A^*-search is admissible.*

3. Multi-agent issues in t-DeLP-POP for ABN

Understanding a plan. Let $\mathsf{Ag} = \{1, \ldots, n\}$ be a set of agents. Each agent i is initially endowed with a planning domain $\mathbb{M}_i = ((\Psi_i, \Delta_i), A_i, G_i)$. $\mathsf{Plans}(\mathbb{M})$ will denote the plans according to domain $\mathbb{M}$. For an agent j to understand a plan Π for some $\mathbb{M}$ (e.g. $\mathbb{M} = \mathbb{M}_i$) communicated by i, it is enough the components of Π are in $\mathbb{M}_j$: literals at Ψ supporting the plan are in Ψ_j; Δ rules used for argument steps are in Δ_j, $A_\Pi \subseteq A_j$[6].

Relativity of threats w.r.t. beliefs. A plan $\Pi = (A_\Pi, \mathsf{Goals}(\Pi), \mathbb{I}(\Pi))$ is defined without threats, because these are relative to the t-de.l.p. (Ψ_i, Δ_i) of the agent i evaluating this plan. For i to detect a threat $\mathcal{B}$ it suffices that $\mathcal{B} \subseteq \Delta_i$. Put the other way round, one can understand a plan without agreeing on the set of threats existing (i.e. on whether Π is a solution for G). Hence, the utility of or preference for a plan can also change due to learning. The set of threats in Π according to $\mathbb{M}$ will be denoted $\mathsf{Threats}^{\mathbb{M}}(\Pi)$.

Commications as expansions of $\mathbb{M}_i$. A dialogue turn will consist in agent i communicating an offer, argument, etc. to an agent j. The information contained (facts, rules, actions) is extracted and learnt by j, by expanding $\mathbb{M}_j$, resp., expanding Ψ_j, Δ_j, A_j (goals are also affected; see offers below). Modeling other agents is done by (instrumental) planning domains of the form $\mathbb{M}_{ji}$, used by j to reason (and exploit) i's difficulties solving G_i (from the point of view of j's own know-how). Since agents' planning domains vary with time, $\mathbb{M}_i$ will be added a superscript denoting the expanded domain at some turn.

Negotiation roles; concession functions. An agent j's service to enforce ℓ, for some customer i's plan Π_i, is an action $\alpha_j \in A_j$ such that $\ell \in \mathsf{X}(\alpha_i) \cap (\mathsf{Goals}(\Pi_i) \cup \mathsf{Threats}^{\mathbb{M}_i}(\Pi_i))$, or an argument $\mathcal{A}$ with $\mathsf{concl}(\mathcal{A}) = \ell$ (the latter at null cost) for similar ℓ. Within a dialogue taking place between agents i and j unfolds, a sub-dialogue between j and k may be triggered ultimately motivated by i's goals.

Since our focus here is in the argumentation aspects of ABN, we will just assume agents are endowed with concession functions: the input of agent i's concession function

[6] An action, say, $\alpha_i \in A_j$ with $i \neq j$ is merely informative. An offer received, or an agreement about α_i, say an exchange of α_i for β_j, will be represented as a new action $(\alpha_i \otimes \beta_j)_j$ *of agent* j in A_j.

w.r.t. an agent j is the set of rival offers addressing the same goal. The output is a new offer improving rival offers that, resp., w.r.t. j's interests, or i's interests.

We might assume as well each agent has a *communication policy* that regulates when one's information (relevant in the dialogue) is not sent, to preserve one's interests. In this paper, though, we assume agents are honest, so all relevant information will be sent.

Social Relations. Finally, we note how social relations between negotiating agents may alter the development of a negotiation dialogue. Among basic social relations in $\mathsf{Ag} \times \mathsf{Ag}$, we consider: cooperative $\equiv_{\mathsf{co}}$, equitable $\simeq_{\mathsf{eq}}$ and hierarchical $\prec_{\mathsf{hi}}$. We simplify by assuming these to be, resp., equivalence relations ($\equiv_{\mathsf{co}}$, $\simeq_{\mathsf{eq}}$) and a partial order $\prec_{\mathsf{hi}}$ in Ag. Agents $a \equiv_{\mathsf{co}} b \equiv_{\mathsf{co}} \ldots$ are cooperative if they share goals $G_a = G_b = \ldots$ (i.e. they adopt each other goals, if consistent); a unique groupwise dialogue suffices for $\equiv_{\mathsf{co}}$ (see [10]). Equitable agents $a \simeq_{\mathsf{eq}} b$ are free to withdraw from a negotiation and need not justify the absence of offers. In contrast, within a hierarchy, say with agents $b \prec_{\mathsf{hi}} a$, the power of a consists in a set of (tacit) agent threats to b, which (by law) cannot be replied or counter-argued. This relation $\prec_{\mathsf{hi}}$ demands a new speech act, b *challenging* a, that consists showing the demands cannot be met, by disclosing to a one's actions/knowledge. (To see this, a must fail to find a plan using this information.)

Speech acts. We define next the speech acts listed in Section 1. These divide into: proposals (offers, demands, acceptance), theoretical arguments (plan threats, agent threats) and practical arguments (persuasion, challenge):

Offers	I propose that I do α if you do β
Demands	I have goal g in plan Π; can you help?
Plan threats	Your/his offer conflicts with this part of my/your plan
Agent threats	If you permit/cause ℓ to occur, I swear I will do α
Persuasion	This agent offers me so-and-so, can you match this offer?
Challenge	In this plan, your demands cannot be met under such deadline, or at all

Argument steps for an agent i's plan Π_i or argument threats to Π_i are generated as in the single-agent case but for Π as a plan in the "planning domain" $(\Psi_j, \Delta_j, A_j, G_i)$.

Definition 1 *Let* $\{x, y\} \subseteq \mathsf{Ag}$*, with plans* Π_x, Π_y*. An* offer κ *from* x *to* y *is:*

- *an arg. step* $\kappa = \mathcal{A} \subseteq \Delta_x$*. This updates* $\mathbb{M}_y$ *as:* $\mathbb{M}_y = ((\Psi_y, \Delta_y \cup \mathcal{A}), \ldots)$.
- *some service exchange* $\kappa = (\alpha_x \otimes \alpha_y, \mathbb{I}(\alpha_x), \mathbb{I}(\alpha_y))$*. Read the old* $\mathbb{M}_y$ *as: "the offer is rejected by* y*"; and model its acceptance in a new* $\mathbb{M}'_y = (\ldots, A_y \cup \{\kappa\}, G_y \cup \{\alpha_y\text{'}\mathsf{ed}\})$*, where* $(\alpha_x \otimes \alpha_y) = (\emptyset, \mathsf{X}(\alpha_x), 0)$.
- *some purchase offer* $\kappa = (\alpha_x \otimes n, \mathbb{I}(\alpha_x))$*. We just expand* $A_y \in \mathbb{M}_j$ *with* $\kappa = (\alpha_x \otimes n) = (\emptyset, \mathsf{X}(\alpha_x), n)$*; (and similarly for* $\mathbb{M}_{x,y}$*). Or,*
- *some purchase offer* $(\alpha_y \otimes n, \mathbb{I}(\alpha_y))$*. Create* $\mathbb{M}'_y = (\ldots, A_y \cup \{(\alpha_y \otimes n)\}, G_y \cup \{\alpha_y\text{'}\mathsf{ed}\})$ *as above, with* $(\alpha_y \otimes n) = (\mathsf{P}(\alpha_y), \mathsf{X}(\alpha_y), \mathsf{cost}(\alpha_y) - n)$.

where $\mathbb{I}(\alpha_x)$ *(resp.* $\mathbb{I}(\alpha_y)$*) is a set of constraints* $\mathbf{t}(\alpha_x) \ \leq / \geq \ m$ *derived from those for action* α_x *in* $\mathbb{I}(\Pi_x(\alpha_x \otimes \alpha_y)$ *(resp. for action* $(\alpha_x \otimes \alpha_y)$*). Here* $m = \Sigma_{\mathbf{t}(\kappa) \leq \mathbf{t}(\kappa') \leq \mathbf{t}(\alpha_{G_x})} \|\kappa'\|$*. All changes to* $\mathbb{M}_y$ *in this definition apply to* $\mathbb{M}_{xy}$ *as well.*

Thus, asking x to add a new constraint $\mathbf{t}(\alpha_x)$ to a previous offer $(\alpha_x \otimes n)$ is a new negotiation and can end up in a different agreement $(\alpha_x \otimes n')$, reflecting x's opportunity cost for α_x under the new temporal constraints $\mathbb{I}(\Pi_x) \leftarrow \mathbb{I}(\Pi_x) \cup \mathbb{I}(\alpha_x)$.

Offers to be sent are generated according to one's *concession policy*. Similarly for threats (see below) which aim to modify the consequences of rival plans, and hence the agent's evaluation of these plans (i.e. her preference relation among plans).

Definition 2 *Let agents $i, j \in \mathsf{Ag}$ with $\mathsf{Goals}(\Pi_i) \neq \emptyset$, for some Π_i current plan in $\mathsf{Plans}(\mathbb{M}_i)$). Components of the planning domain $\mathbb{M}_x$ are expanded as follows:*

demand$_{i \triangleright j}$: Π_i	*(j)* $\Pi_i \in \mathsf{Plans}(\mathbb{M}_{ji})$
offer$_{j \triangleright i}$: $\Pi_i(\kappa)$	*(i)* $\kappa \subseteq \Delta_i$ *or* $(\alpha_i \otimes \beta_j) \in A_i$ *(similarly for* $\mathbb{M}_{ji}$*)*
plan threat$_{j \triangleright i}$: $\mathcal{B}$	*(i)* $\Delta_i \cup \mathcal{B}$
agent threat$_{j \triangleright i}$: (ℓ, α)	*(i)* $\Delta_i \cup \{\ell' \prec (\ell, \lVert\alpha\rVert) \mid \ell' \in \mathsf{X}(\alpha)\}$.
	(j) $G_j \cup \{p^\star\}$*, for the new literal* $p^\star := \mathrm{satisfied}\text{-}(\ell, \alpha)$ *, with*
	(j) $\{p^\star \prec {\sim}(\ell, 0); p^\star \prec (\ell, 0), \mathsf{X}(\alpha) \times \{0\}\} \subseteq \Delta_j$
persuasion$_{i \triangleright j}$: $\Pi_i(\kappa_k)$	*(j)* $\Pi_i(\kappa_k) \in \mathsf{Plans}(\mathbb{M}_{ji})$.
challenge$_{j \triangleright i}$: Π_i	$\mathbb{M}_j = (\Psi_j \cup \Psi_i^*, \Delta_j \cup \Delta_i^*, A_j \cup A_i^*, G_j)$.

where the relevant subsets Ψ_i^*, Δ_i^*, A_i^* (of $\mathbb{M}_i$) are defined as in the dialogues of [10].

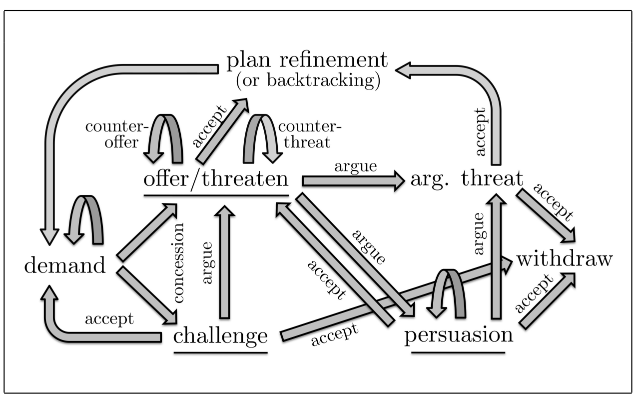

Figure 2. The ABN protocol. Underlined moves may involve nested dialogues with new agents.

Note that a threatened agent i will make public the threats received by j to the other parties to point out these new existing "argument threats" in their suggested plans.
Practical preference. During plan search two notions of preference play a role. For plan refinement, the best plan is selected by each agent i, solely based on its cost $\mathsf{cost}(\Pi) := \Sigma_{\alpha \in A_\Pi} \mathsf{cost}(\alpha)$. The remaining plans can still be "improved" by persuading other agents by showing the existence of better rival offers. In this persuasion moves, the preference criterion has stronger conditions, namely, $\Pi \succeq_i^{\mathbb{M}_{ji}} \Pi'$ if and only if

$$\mathsf{Threats}^{\mathbb{M}_{ji}}(\Pi) \subseteq \mathsf{Threats}^{\mathbb{M}_{ji}}(\Pi'), \mathsf{Goals}(\Pi) \subseteq \mathsf{Goals}(\Pi') \text{ and } \mathsf{cost}(\Pi) \leq \mathsf{cost}(\Pi)$$

with one of these inequalities being strict: resp., $\subsetneq$ or $\subsetneq$ or $<$. Here $\mathbb{M}_{ji}$ denotes $\mathbb{M}_i$ if $j = i$. Note this preference criterion is subject to discussion because of the relativity of threats. Hence, a persuasion move can be replied by a plan threat, which can be replied by a persuasion move from a refined plan, etc. (see Figure 2).

The preference $\succeq_i^{\mathbb{M}_{ji}}$ restricts j's counter-offers to i: given two rival plans $\Pi_i \succeq_i^{\mathbb{M}_{ji}} \Pi'_i$, involving offers with different agents and Π'_i containing an offer $(\alpha_i \otimes \beta_j)$, new offers from j, say $(\alpha'_i \otimes \beta'_j)$, must make the new plan Π''_i improve Π'_i: $\Pi''_i \succeq_i^{\mathbb{M}_{ji}} \Pi'_i$. In the particular case where $\Pi_i = \Pi^\star_i(\dots)$ and $\Pi'_i = \Pi^\star(\alpha_i \otimes \beta_j)$, new offers Π''_i must be competitive: $\Pi''_i \succeq_i^{\mathbb{M}_{ji}} \Pi_i$.

4. A protocol for ABN dialogues.

Even if dialogues are assumed to occur concurrently, we model them as taking place in a sequential way. Thus, turns n encode a pair $f : n \mapsto \langle i, j\rangle$ of agents i and j, resp. the sender at and the receiver of information at turn n; here $f : \mathbb{N} \to \mathsf{Ag}^2$ is an enumeration of Ag^2 always following the same order, i.e. satisfying $f(n) = f(n + |\mathsf{Ag}| \cdot (|\mathsf{Ag}| - 1))$. The content sent at a turn of the form $\langle i, j\rangle$ is fully determined by the current $\mathbb{M}_i$, $\mathbb{M}_{ij}$ domains, and previous (rival) offers in turns of the form $\langle i, j\rangle$ or $\langle k, i\rangle$ (where $k \neq i, j$).

Definition 3 *Given turn $n \xrightarrow{f} \langle i, j\rangle$, the messages sent by i, and the changes in j are:*

- *Each message of i at n follows the next-step(s) in Figure 2, for each message in the previous turn of the form $\langle j, i\rangle$. Agent i updates $\mathbb{M}_{ij}$, similarly to that of $\mathbb{M}_j$, see next.*
- *The components of each domain $\mathbb{M}_j, \mathbb{M}'_j$, etc are updated by each message received by i as usual. Agent j updates each set $\mathsf{Plans}(\mathbb{M}_j)$, $\mathsf{Plans}(\mathbb{M}'_j)$, etc. with new refine–ment steps available, as well as (re-)evaluating $\mathsf{Threats}^{\mathbb{M}_j}(\Pi)$, etc. for each (explored) plan Π in old $\mathsf{Plans}(\mathbb{M}_j)$, etc. A new $\mathbb{M}'_j$ is created for each new offer received.*

Proposition 1 *Given a sequence $(\mathbb{M}_i)_{i\in\mathsf{Ag}}$, and concession functions γ_i, the ABN dialogues terminate in a finite time.*

(Proof Sketch) The reason is that: since $\mathsf{Ag}, \mathsf{Var}$ are finite and each $\mathbf{t}(\alpha_{G_x})$ is bounded, each possible set Ψ_x, Δ_x, A_x and G_x is finite, and so is the number of $\mathbb{M}_x$'s for each $x \in \mathsf{Ag}$. Also, the set of offers is finite. Hence, the refinements of any Π must be finite, and the length of $\Pi \in \mathsf{Plans}(\mathbb{M})$ must be bounded. Thus, each set $\mathsf{Plans}(\mathbb{M}_x)$ is finite. Thus, the negotiation space is finite and so is any dialogue between x and y.

Once a dialogue terminates, which agreements are reached depends on whether we take offers $(\alpha_i \otimes \beta_j)$ sent by i as committing i (if accepted by j) or not. If they do, the previous ABN protocol determines which agreements are reached. If offers do not commit agents that proposed them, the agreement problem turns out to be that of forming overlapping coalitions in the line of [1] (i.e. a problem in cooperative game-theory).

4.1. Case study: negotiating a deadline.

In the following, we assume action $\{\mathsf{have_x(o)}, \sim\!\mathsf{have_y(o)}\} \xleftarrow{\mathsf{give_{y\triangleright x}(o)}} \{\mathsf{have_y(o)}\}$ is known by all agents.

Example 1 *Given a company with manager* a *and a vetting agent* b *such that* $\mathsf{b} \prec_{\mathsf{hi}} \mathsf{a}$, *let* $\mathsf{c_0}$ *be a customer asking* a *to change their current deal. (1)* a *orders (i.e. demands)* b *to vet all the customers* $\mathsf{c_0}, \dots, \mathsf{c_n}$ *in the area, in 8 hrs. (2)* b *argues this cannot be done (i.e.* b *challenges* a*). As a reply, the manager* a *can send two new demands: (3) to have* *just the important customer c_0 vetted; or (3') to keep the goals but extend the deadline to tomorrow morning, i.e. in 24 hrs. Agent* b *sends solution plans, resp., at lines (4) and (4'). In (4')* b *replies by asking to be paid for the extra work (at one coin per extra hour).*

---

(1) demand$_{\mathsf{a}\rhd\mathsf{b}}$	$\Pi_{\mathsf{a}} = (\emptyset, \{\mathsf{have_a}(\mathsf{docu}(\mathsf{c_k}))\}_{\mathsf{k}\leq\mathsf{n}}, \{\mathbf{t}(\alpha_{\mathsf{G_a}}) \leq 8\})$
(2) challenge$_{\mathsf{b}\rhd\mathsf{a}}$	$A^*_{\mathsf{b}} = \{\mathsf{have_b}(\mathsf{docu}(\mathsf{c_k})\} \xleftarrow{\mathsf{vet_b}(\mathsf{c_k})} \}$, with $\|\cdot\| = 3\}$
	$\Psi^*_{\mathsf{b}} = \{\sim\mathsf{have_b}(\mathsf{docu}(\mathsf{c_k}))\}_{\mathsf{k}\leq\mathsf{n}}\}$
	a *checks* $\mathbb{M}^*_{\mathsf{a}} = ((\Psi_a \cup \Psi^*_b, \Delta_a), A_a \cup A^*_b, G_a)$ *has no solution.*
(3) demand$_{\mathsf{a}\rhd\mathsf{b}}$	$\Pi_{\mathsf{a}} = (\emptyset, \{\mathsf{have_j}(\mathsf{docu}(\mathsf{c_0}))\}, \{\mathbf{t}(\alpha_{\mathsf{G_a}}) \leq 8\})$
(4) offer$_{b\rhd a}$	$\Pi_{\mathsf{a}}(\{\mathsf{give_{b\rhd a}}(\mathsf{docu}(\mathsf{c_0}))), \mathsf{vet_b}(\mathsf{c_0})\}, \emptyset, \{\mathbf{t}(\mathsf{vet_b}(\mathsf{c_0})) + 3 < 8\})$
(3') demand$_{\mathsf{a}\rhd\mathsf{b}}$	$\Pi_{\mathsf{a}} = (\emptyset, \{\mathsf{have_a}(\mathsf{docu}(\mathsf{c_k}))\}_{\mathsf{k}\leq\mathsf{n}}, \{\mathbf{t}(\alpha_{\mathsf{G_a}}) \leq 24\})$
(4') offer$_{\mathsf{b}\rhd\mathsf{a}}$	$\Pi_{\mathsf{a}} = (\{(\mathsf{give_{b\rhd a}}(\mathsf{docu}(\mathsf{c_k})) \otimes 3), \mathsf{vet_b}(\mathsf{c_k})\}_{\mathsf{k}\leq\mathsf{n}}, \emptyset,$
	$\{\mathbf{t}(\mathsf{vet_b}(\mathsf{c_k})) + 3 \leq \mathbf{t}(\mathsf{vet_b}(\mathsf{c_{k+1}})) \leq 24\}_{\mathsf{k}\leq\mathsf{n}})$

4.2. Case study: hanging picture and mirror.

Example 2 *(Adapted from [14]) Agent* a*, endowed with a nail and a picture, wants to hang the picture:* $\mathsf{picture.in.wall^a} \in G_{\mathsf{a}}$*; while b, endowed with a hammer and a mirror, wants to hang the mirror while keeping the hammer:* $\mathsf{mirror.in.wall^b}, \mathsf{have_b}(\mathsf{hammer}) \in G_{\mathsf{b}}$*. Both agents' deadlines are set to 4 hrs. Agents' actions* A_{x} *are defined for any* $\mathsf{x}, \mathsf{u} \in \{\mathsf{a}, \mathsf{b}\}$*, where* x *denotes the executing agent, and* u *the proprietor of* $\mathsf{wall^u}$*:*

$$\begin{array}{rcl}
\{\mathsf{nail.in.wall^u_x}\} & \xleftarrow{\mathsf{hammer}(\mathsf{nail.wall^u})_{\mathsf{x}}} & \{\mathsf{have_x}(\mathsf{hammer}), \mathsf{have_x}(\mathsf{nail})\} \\
\{\mathsf{picture.in.wall^u_x}\} & \xleftarrow{\mathsf{hang}(\mathsf{picture.wall^u})_{\mathsf{x}}} & \{\mathsf{nail.in.wall^u}, \mathsf{have_x}(\mathsf{picture})\} \\
\{\mathsf{mirror.in.wall^u_x}\} & \xleftarrow{\mathsf{hang}(\mathsf{mirror^u})_{\mathsf{x}}} & \{\mathsf{nail.in.wall^u}, \mathsf{have_x}(\mathsf{mirror})\}
\end{array}$$

No agent's goal g *is assumed to persist* $\delta_g \notin \Delta_{\mathsf{u}}$*, since persistence of* g *is dependent on that of* $\mathsf{nail.in.wall^u}$*. In fact,* b *is sick, so she cannot hammer strong enough. This is modeled by rules in* $\Delta_{\mathsf{a}}, \Delta_{\mathsf{b}}$*:* $\langle \mathsf{weak.nail.wall^u_b}, \rangle \prec \langle \mathsf{nail.in.wall^u_b}, \mathsf{t}\rangle, \langle \mathsf{sick_b}, \mathsf{t}\rangle$ *and* $\langle \sim\mathsf{nail.in.wall^u}, \mathsf{t}+2\rangle \prec \langle \mathsf{nail.in.wall^u}, \mathsf{weak.nail.in.wall^u_b}, \mathsf{t}\rangle, \langle \mathsf{weak.nail.wall^u}, \mathsf{t}\rangle$.

After initial demands, a *offers* $\mathsf{give_{a\rhd b}}(\mathsf{nail}) \otimes (\mathsf{give_{b\rhd a}}(\mathsf{hammer}) \otimes \mathsf{give_{a\rhd b}}(\mathsf{hammer}))$ *with obvious constraints. Then* b *challenges the resulting plan by communicating he is sick; agent* a *is able to detect a threat in the proposed plan to* b*. Then,* b *proposes a new offer* $\mathsf{hammer}(\mathsf{nail.wall^b})_{\mathsf{a}} \otimes (\mathsf{give_{b\rhd a}}(\mathsf{hammer}) \otimes \mathsf{give_{a\rhd b}}(\mathsf{hammer}))$.

4.3. Case study: threats in a 2*-vendors* 1*-customer scenario.*

Example 3 *The manager* a *of a building company learns a container of bricks has been stolen from a building. She phones (among others) agent* b *from a security company, who offers to protect the building for* m *coins per night (the action* $\mathsf{protect_c}$ *at* t *does preserve all* $\langle \mathsf{have_a}(\cdot)$ *literals up to hour* $t+8$*). Agent* a *replies by producing a better offer at cost* m' *from another agent* c*. Then,* b *suggests that some robberies could affect* again *the construction schedule for this building. The rule* $\langle \sim\mathsf{have_b}(\mathsf{bricks}), 24\rangle \prec \langle \sim\mathsf{protect_c}\text{'}\mathsf{ed}, 0\rangle, \langle \mathsf{protect_b}\text{'}\mathsf{ed}, 0\rangle$ *is learnt by the manager. Since* $\mathsf{cost}(\mathsf{buy}(\mathsf{bricks})) + \mathsf{m}' > \mathsf{m}$*, agent* a *may decide to pay* m *to* b*. Or, if a po-*

liceman p *exists,* a *can reply by (counter-) threatening* b *to inform* p*. Since the rule* $\mathsf{jailed_b} \ \prec (\mathsf{inform_{a \triangleright p}(criminal(b))}, 4)$ *is in both sets* Δ_{ab} *and* Δ_{b}*, agent* c *cancels her own threat, so* b *deletes the corresponding rule, and* c *withdraws from the dialogue.*

5. Conclusions and Future work.

We have presented a protocol for negotiation dialogues that take place within a society of planner-reasoner agents. We showed some basic properties, and studied the dialogues for several ABN scenarios extracted from the literature. These case studies show the proposed framework for the agent architecture is quite expressive for ABN scenarios. As for future work, we would like to study the relation between algorithmic search and game-theoretic properties of agreements in the proposed ABN protocol: e.g. Pareto-optimality of the sequence of plans agreed upon, etc.

References

[1] M. Albizuri, J. Aurrecoechea, J. Zarzuelo *Configuration values: Extensions of the coalitional Owen value*, Games and Economic Behavior 57, pp. 1–17 (2006)

[2] L. Amgoud, Y. Dimmopoulos, and P. MoraitisÊ*A Unified and General Framework for Argumentation-based Negotiation* Proc. of AAMAS 2007

[3] L. Amgoud and H. Prade *Formal Handling of Threats and Rewards in a Negotiation Dialogue*, Proc of ArgMAS 2005

[4] A. Artikis, M. Sergot and J. Pitt *An executable specification of a formal argumentation protocol* Artificial Intelligence 171, pp. 776Ð804 (2007)

[5] Y. Dimopoulos, P. Moraitis, L. Amgoud. *Characterizing the Outcomes of Argumentation-Based Integrative Negotiation*, Proc. of IAT 2008, 456-460.

[6] A. García and G. Simari. Defeasible logic programming: An argumentative approach. *Theory and Practice of Logic Programming*, 4:95–138, 2004.

[7] D. García, A. García, and G. Simari. Defeasible reasoning and partial order planning. In Proc. of the *5th International Conference on Foundations of information and knowledge systems, FoIKS 2008*, LNCS 4932, pp. 311–328, 2008.

[8] J. Grant, S. Kraus and D. Perlis *A logic-based model of intention formation and action for multi-agent subcontracting*, Artificial Intelligence, 163. 163–201, 2005.

[9] P. Pardo *Temporal defeasible logic-based planning t-DeLP-POP for physically correct plans.* Available at http://www.iiia.csic.es/files/pdfs/Pardo-GenPlanAAAI11.pdf.

[10] P. Pardo. S. Pajares, E. Onaindía, L. Godo, P. Dellunde. *Multi-agent argumentation for cooperative planning in DeLP-POP*, Proc. of AAMAS 2011, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2-6, 2011, Taipei, Taiwan, pp. 971-978.

[11] P. Pardo and Lluís Godo *t-DeLP: a temporal extension of the defeasible logic programming argumentative framework* (Submitted). Available at http://www.iiia.csic.es/files/pdfs/tDeLP-submitted.pdf

[12] J. Pearl. *Heuristics: Intelligent Search Strategies for Computer Problem Solving*. Addison-Wesley, 1984.

[13] J. Penberthy and D. Weld. Ucpop: A sound, complete, partial order planner for adl. In Proc. of the *3rd International Conference on Knowledge Representation and Reasoning (KR'92)*, pp. 103–114, 1992.

[14] S. Parsons, C. Sierra and N. Jennings *Agents that reason and negotiate by arguing*. Journal of Logic and Computation 8 (3), 261-292, 1998.

[15] C. Sierra, N. Jennings, P. Noriega and S. Parsons *A framework for argumentation-based negotiation*, In Proceedings of the 4th International Workshop on Intelligent Agents IV, Agent Theories, Architectures, and Languages (ATAL'97), 177 - 192, 1997.

[16] K. Sycara, S. Kraus and A. Evechnik *Reaching agreements through Argumentation: A Logical Model and Implementation*, Artificial Intelligence, 104 (1-2), pp. 1-69 (1998)