Assumption-Based Argumentation Dialogues

Xiuyi Fan and Francesca Toni
Department of Computing, Imperial College London, UK
{xf309,ft}@imperial.ac.uk

Abstract

We propose a formal model for argumentation-based dialogues between agents, using assumption-based argumentation (ABA). The model is given in terms of ABA-specific utterances, trees drawn from dialogues and legal-move and outcome functions. We prove a formal connection between these dialogues and argumentation semantics. We illustrate persuasion as an application of the dialogue model.

1 Introduction

Argumentation-based agent dialogues have been widely studied, e.g. in [McBurney and Parsons, 2009; Black and Hunter, 2009; Prakken, 2005; 2006], as a framework to support agreement amongst artificially intelligent agents. We present a novel formal model of argumentation-based dialogues between (two) agents, using *assumption-based argumentation* (ABA) [Bondarenko *et al.*, 1997; Dung *et al.*, 2009]. This is a general-purpose, widely applicable form of argumentation where arguments are built from *rules* and supported by *assumptions*, and attacks against arguments are directed at the assumptions supporting the arguments, and are provided by arguments for the *contrary* of assumptions.

In our dialogue model, agents can utter claims (to be debated), rules, assumptions and contraries. Thus, dialogues "build" an ABA framework shared between the agents.

The model is given in terms of (various kinds of) legal-move functions, to determine which moves agents can make during dialogues, and outcome functions, to determine whether dialogues have been successful. These functions are defined in terms of *dialectical trees* drawn from the dialogues (and implicitly constructed during them).

To prove soundness of our approach, we connect our dialogue model with the admissibility semantics for ABA, sanctioning a set of assumptions/arguments as *admissible* iff it does not attack itself and it counter-attacks all attacks against it. In particular, we prove that the claim debated by a successful dialogue is supported by a set of assumptions/argument in an admissible set (and this is the set of assumptions/arguments identified during the dialogue) w.r.t. the shared ABA framework drawn from the dialogue. This result relies upon a correspondence result between dialectical trees and the concrete dispute trees given in [Dung *et al.*, 2006] to pave the way towards a computational counterpart (in the form of dispute derivations) for the admissibility semantics.

Our dialogue model is generic in that it does not focus on any particular dialogue type, e.g. information seeking, persuasion or negotiation. However, for the sake of illustration, we demonstrate persuasion as an application of our model.

The paper is organised as follows. Section 2 reviews the ABA framework. Section 3 presents the dialogue model. Section 4 introduces dialectical trees. Section 5 gives our main formal results. Section 6 demonstrates a persuasion dialogue. Section 7 discusses related works. Section 8 concludes.

2 Background

An ABA framework [Bondarenko *et al.*, 1997; Dung *et al.*, 2009] is a tuple $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \mathcal{C} \rangle$ where

- $\langle \mathcal{L}, \mathcal{R} \rangle$ is a deductive system, with $\mathcal{L}$ the *language* and $\mathcal{R}$ a set of *rules* of the form $s_0 \leftarrow s_1, \ldots, s_m (m \geq 0)$;
- $\mathcal{A} \subseteq \mathcal{L}$ is a (non-empty) set, referred to as *assumptions*;
- $\mathcal{C}$ is a total mapping from $\mathcal{A}$ into $\mathcal{L}$, where $\mathcal{C}(a)$ is referred to as the *contrary* of a.

Given a rule ρ of the form $s_0 \leftarrow s_1, \ldots, s_m$, s_0 is referred as the *head* and $s_1, \ldots, s_m$ as the *body*, $Head(\rho) = s_0$ and $Body(\rho) = \{s_1, \ldots, s_m\}$. As in [Dung *et al.*, 2006], we enforce that ABA frameworks are *flat*, i.e. assumptions do not occur in the head of rules. Also, without loss of generality, we enforce that no two assumptions may have the same contrary.

In ABA, *arguments* are deductions of claims using rules and supported by sets of assumptions, and *attacks* are directed at the assumptions in the support of arguments. Informally, following [Dung *et al.*, 2009]:

- *an argument for (the claim)* $c \in \mathcal{L}$ *supported by* $S \subseteq \mathcal{A}$ ($S \vdash c$ in short) is a tree with nodes labelled by sentences in $\mathcal{L}$ or by τ,[1] the root labelled by c, leaves either τ or assumptions in S, and non-leaves s with as children the elements of the body of some rule with head s;
- *an argument* $S_1 \vdash c_1$ *attacks an argument* $S_2 \vdash c_2$ iff c_1 is the contrary of one of the assumptions in S_2.

Attacks between (sets of) arguments correspond in ABA to attacks between sets of assumptions:

[1] $\tau \notin \mathcal{L}$ represents "true" and stands for the empty body of rules.

- *a set of assumptions A attacks a set of assumptions B* iff an argument supported by a subset of A attacks an argument supported by a subset of B.

With argument and attack defined, standard argumentation semantics can be applied in ABA [Dung *et al.*, 2009]. We focus on the admissibility semantics, as follows:

- *a set of assumptions is admissible* iff it does not attack itself and it attacks all sets of assumptions that attack it;
- *an argument $S \vdash s$ belongs to an admissible extension supported by* $\Delta \subseteq \mathcal{A}$ iff $S \subseteq \Delta$ and Δ is admissible.

Our main result will be proven using the concrete dispute trees of [Dung *et al.*, 2006], where a *concrete dispute tree* for a sentence $s \in \mathcal{L}$ is a (possibly infinite) tree $\mathcal{T}^c$ such that

1. every node of $\mathcal{T}^c$ is labelled by a multi-set of sentences in $\mathcal{L}$ and is either a *proponent* node or an *opponent* node;
2. the root of $\mathcal{T}^c$ is a proponent node labelled by $\{s\}$;
3. given a proponent node N labelled by $\mathcal{P}$, if $\mathcal{P}$ is empty, then N is a leaf; otherwise, there exists some selected[2] occurrence of a sentence p in $\mathcal{P}$ and
   - if $p \in \mathcal{A}$, then there is one child of N, which is an opponent node labelled by $\{\mathcal{C}(p)\}$, and one child of N that is a proponent node labelled by $\mathcal{P} - \{p\}$;
   - If $p \notin \mathcal{A}$, then there is some $\rho \in \mathcal{R}$ with $Body(\rho) = S$ and there is exactly one child of N, which is a proponent node labelled by $\mathcal{P} - \{p\} \cup S$;
4. given an opponent node N labelled by $\mathcal{O}$, then $\mathcal{O} \neq \{\}$, there is a selected occurrence of a sentence o in $\mathcal{O}$ and
   (i) if $o \in \mathcal{A}$, then (a) either o is *ignored* and there is exactly one opponent child of N, labelled by $\mathcal{O}$–$\{o\}$; (b) or o is chosen as *culprit*, and there is exactly one proponent child of N, labelled by $\{\mathcal{C}(o)\}$;
   (ii) If $o \notin \mathcal{A}$ and there is no $\rho \in \mathcal{R}$ with $Head(\rho) = o$, then N is a leaf; else, for every $o \leftarrow S \in \mathcal{R}$, there is an opponent child of N labelled by $\mathcal{O} - \{o\} \cup S$;
5. there is no infinite sequence of proponent nodes in $\mathcal{T}^c$; there are no nodes in $\mathcal{T}^c$ except those given by 1-4.

Given a concrete dispute tree $\mathcal{T}^c$, the *defence set* of $\mathcal{T}^c$ ($\mathcal{D}ef(\mathcal{T}^c)$) is the union of all assumptions in proponent nodes and the *culprits* of $\mathcal{T}^c$ ($\mathcal{C}ul(\mathcal{T}^c)$) are all assumptions chosen as culprits in 4(i)(b). A concrete dispute tree is *admissible* iff the intersection of its defence set and its culprits is empty. The defence set of an admissible concrete dispute tree $\mathcal{T}^c$ for a sentence s is admissible and there exists an argument $S \vdash s$ that belongs to an admissible extension supported by the defence set of $\mathcal{T}^c$ (corollary 6.1 in [Dung *et al.*, 2006]).

3 Dialogues

We define dialogues as sequences of utterances between two agents a_1 and a_2 sharing a common language $\mathcal{L}$. Formally:

Definition 1. An *utterance* from agent a_i to agent a_j ($i, j = 1, 2$, $i \neq j$) is a tuple $\langle a_i, a_j, InReply, C, ID\rangle$, where:

- C (the *content*) is of one of the following forms:
  $clm(s)$ for some $s \in \mathcal{L}$ (a *claim*),
  $rl(s_0 \leftarrow s_1, \ldots, s_m)$ for some $s_0, \ldots, s_m \in \mathcal{L}$ (a *rule*),
  $asm(a)$ for some $a \in \mathcal{L}$ (an *assumption*),
  $ctr(a, s)$ for some $a, s \in \mathcal{L}$ (a *contrary*),
  a *pass* sentence π, such that $\pi \notin \mathcal{L}$.
- $ID \in \mathbb{N}$ (the *identifier*).
- $InReply \in \mathbb{N} \cup \{0\}$ (the *target*); $InReply < ID$.

We refer to an utterance with content π (other than π) as a *pass-utterance* (*non-pass-utterance* resp.).

Intuitively, a pass indicates that the agent does not have or want to contribute information at that point in the dialogue.

Definition 2. For any two utterances $u_i \neq u_j$, u_j is *related to* u_i iff $u_i = \langle \_, \_, \_, C_i, ID\rangle$, $u_j = \langle \_, \_, ID, C_j, \_\rangle$,[3] and one of the following cases holds:

1. $C_j = rl(\rho_j)$, $Head(\rho_j) = h$ and either $C_i = rl(\rho_i)$ with $h \in Body(\rho_i)$, or $C_i = ctr(\_, h)$, or $C_i = clm(h)$;
2. $C_j = asm(a)$ and either $C_i = rl(\rho)$ with $a \in Body(\rho)$, or $C_i = ctr(\_, a)$, or $C_i = clm(a)$;
3. $C_j = ctr(a, \_)$ and $C_i = asm(a)$.

Intuitively, an utterance is related to another if it contributes to expanding an argument (case 1), identifies an assumption in the support of an argument (case 2) or starts the construction of a counter-argument (case 3). Note that an utterance may be related to an utterance from the same agent or not. Also, no pass-utterance can be related to an utterance and no utterance can be related to a pass-utterance.

Definition 3. A *dialogue $D^{a_i}_{a_j}(s)$ (between agents a_i and a_j, for claim $s \in \mathcal{L}$)*, $i, j = 1, 2$, $i \neq j$, is a finite sequence $\langle u_1, \ldots, u_n\rangle$, $n \geq 0$, where each u_l, $l = 1, \ldots, n$, is an utterance from a_i when l is odd, and a_j when l is even, and:

1. the content of the first utterance, u_1, is $clm(s)$; no other utterance has $clm(\_)$ as its content;
2. each non-pass-utterance other than the claim utterance is related to its target utterance;
3. the target of pass- and claim utterances is 0;
4. no two consecutive utterances are pass-utterances, other than possibly the last two utterances, u_{n-1} and u_n.

If the last two utterances are pass-utterances, then $D^{a_i}_{a_j}(s)$ is referred as *complete*.

Note that in this definition of dialogue, we implicitly force a strict interleaving between agents.

Below, $\mathcal{U}$ and $\mathcal{D}$ stand for the sets, resp., of all utterances as in definition 1 and of all dialogues as in definition 3.

Example 3.1. A possible (complete) dialogue $D^{a_1}_{a_2}(k)$ is:[4]

a_1	a_2
$\langle a_1, a_2, 0, clm(k), 1\rangle$	$\langle a_2, a_1, 0, \pi, 2\rangle$
$\langle a_1, a_2, 1, rl(k \leftarrow a), 3\rangle$	$\langle a_2, a_1, 3, asm(a), 4\rangle$
$\langle a_1, a_2, 4, ctr(a, q), 5\rangle$	$\langle a_2, a_1, 5, rl(q \leftarrow b), 6\rangle$
$\langle a_1, a_2, 6, asm(b), 7\rangle$	$\langle a_2, a_1, 7, ctr(b, c), 8\rangle$
$\langle a_1, a_2, 8, asm(c), 9\rangle$	$\langle a_2, a_1, 9, ctr(c, r), 10\rangle$
$\langle a_1, a_2, 0, \pi, 11\rangle$	$\langle a_2, a_1, 0, \pi, 12\rangle$

[2]Selections are performed by a given *selection function*, that picks a sentence in an input multi-set [Dung *et al.*, 2006].

[3]Throughout, _ stands for an an anonymous variable as in Prolog.

[4]Here and from now on identifiers of utterances coincide with their position in dialogues.

By means of dialogues agents exchange information and build a shared framework, as follows:

Definition 4. The *framework drawn from a dialogue* $\delta = \langle u_1, \ldots, u_n \rangle$ is $\langle \mathcal{L}, \mathcal{R}_\delta, \mathcal{A}_\delta, \mathcal{C}_\delta \rangle$ where[5]

- $\mathcal{R}_\delta = \{r | rl(\rho)$ is the content of some u_i in $\delta\}$;
- $\mathcal{A}_\delta = \{a | asm(a)$ is the content of some u_i in $\delta\}$;
- $\mathcal{C}_\delta$ is a mapping such that, for any $a \in \mathcal{A}_\delta$, $\mathcal{C}_\delta(a) = c$ such that $ctr(a, c)$ is the content of some u_i in δ, if one exists, and is undefined otherwise.

The framework drawn from the dialogue in example 3.1 is $F_\delta = \langle \mathcal{L}, \mathcal{R}_\delta, \mathcal{A}_\delta, \mathcal{C}_\delta \rangle$, in which
$\mathcal{A}_\delta = \{a, b, c\}$;
$\mathcal{R}_\delta = \{k \leftarrow a, q \leftarrow b\}$;
$\mathcal{C}_\delta$ is such that $\mathcal{C}_\delta(a) = q, \mathcal{C}_\delta(b) = c, \mathcal{C}_\delta(c) = r$.

Note that F_δ in this example is a flat ABA framework, but, in general, the framework drawn from a dialogue may not be an ABA framework, since $\mathcal{C}$ may not be total. Even when it is an ABA framework, it may not be flat, as the agents may disagree on the assumptions. Finally, there may be a non-deterministic choice as to what the contrary of an assumption is in $\mathcal{C}_\delta$, as the agents may have uttered different contrary utterances for this assumption. In the remainder of this section, we refine our dialogues so that the frameworks drawn from them are flat ABA frameworks and the choice of contrary is deterministic. This refinement builds upon the notion of *legal-move function* to restrict the kind of utterances allowed in dialogues , similarly to communication protocols.

Definition 5. A *legal-move function* is a mapping $\lambda : \mathcal{D} \mapsto \mathcal{U}$ such that, given $\delta = D^{a_i}_{a_j}(s) = \langle u_1, \ldots, u_n \rangle$:

$$\lambda(\delta) = \begin{cases} \langle a_i, a_j, 0, clm(s), 1 \rangle & \text{if } n = 0 \\ \langle a_x, a_y, t, C, n+1 \rangle & \text{if } n > 0 \end{cases}$$

where $x = i$ ($x = j$) when n is even (odd resp.), $x \neq y$, and

1. if $C \neq \pi$, then there exists no $i, 1 \leq i \leq n$, such that $u_i = \langle \_, \_, t, C, i \rangle$;
2. if $C = ctr(a, c)$ then there exists no $u_i = \langle \_, \_, \_, ctr(a, c'), i \rangle$, for $1 \leq i \leq n$ and $c' \neq c$.

This definition imposes that there is no repeated utterance to the same target in a dialogue (condition 1), and that assumptions have a single contrary (condition 2). Condition 2 guarantees a deterministic choice of contrary. However, the definition of legal-move function does not impose any "mentalistic" requirement on agents, such as that they utter information they hold true within their ABA framework.

Definition 6. A *flat legal-move function* is such that if, for $n \geq 0$, $\lambda(\langle u_1, \ldots, u_n \rangle) = \langle a_x, a_y, t, C, n+1 \rangle$, then

- $C = asm(a)$ only if there exists no u_i for $1 \leq i \leq n$ with content $rl(\rho)$ and $Head(\rho) = a$;
- $C = rl(\rho)$ only if there exists no u_i for $1 \leq i \leq n$ with content $asm(a)$ and $Head(\rho) = a$.

Trivially, the framework drawn from a dialogue generated by a flat legal-move function, if an ABA framework, is flat.

[5]Here and throughout the paper $\mathcal{L}$ is implicit.

Legal-move functions provide some guidance as to what agents should utter during dialogues. In order to guarantee that the contrary mapping in the framework drawn from a dialogue is total, we will use the notion of *outcome function*, checking specific properties in a generated dialogue:

Definition 7. An *outcome function* is a mapping $\omega : \mathcal{D} \mapsto \{true, false\}$. The *ABA outcome function* ω_{ABA} is such that, given $\delta = \langle u_1, \ldots, u_n \rangle \in \mathcal{D}$, $\omega_{ABA}(\delta) = true$ iff $\forall u_i = \langle \_, \_, \_, asm(a), i \rangle \; \exists u_j = \langle \_, \_, i, ctr(a, c), j \rangle$, for $1 \leq i \leq j \leq n$.

We focus on dialogues where each utterance results from applying a flat legal-move function to the dialogue prior to that utterance, and for which ω_{ABA} is true. We refer to these dialogues as *ABA dialogues*. Trivially:

Proposition 1. The framework drawn from an ABA dialogue is a flat ABA framework.

Below we refine the notions of legal-move and outcome functions to guarantee that dialogues compute admissible arguments. These refinements are given using *dialectical trees*.

4 Dialectical trees

Nodes of dialectical trees are either proponent or opponent nodes, as in the case of concrete dispute trees (see section 2). However, in a dialectical tree, nodes are labelled with *pairs* of multi-sets of sentences and are *associated with an utterance* in the dialogue from which the tree is extracted. This extraction ignores the pass-utterances, i.e. a dialectical tree is extracted from the *π-pruned sequence obtained from a dialogue*, consisting of all non-pass utterances in the dialogue. Note that, since no non-pass utterance has a pass-utterance as its target (see definition 3), the target of every utterance in a π-pruned sequence is guaranteed to be in this sequence.

The label of each node in a dialectical tree represents an argument's claim or support (see definition 10 below). The first component of this label holds sentences (referred to as *marked*) that have been "declared" as assumptions in some utterance in the dialogue prior to (the utterance corresponding to) that node. The second component holds sentences (referred to as *unmarked*) that have been introduced in the dialogue but have not yet been "discussed" in any utterance. Formally:

Definition 8. A *dialectical tree* drawn from a dialogue $\delta = D^{a_i}_{a_j}(s) = \langle u_1, \ldots, u_n \rangle$ ($n \geq 0$) is a tree $\mathcal{T}(\delta)$ whose nodes are tuples $([S_{md}, S_{umd}], L, U)$ where

- S_{md} and S_{umd} are (multi-)sets[6] of sentences in $\mathcal{L}$
- L is either P (*proponent*) or O (*opponent*)
- $U \in \mathbb{N}$ (the *u-id*)

and $\mathcal{T}(\delta)$ is $\mathcal{T}^m(\delta)$ in the sequence $\mathcal{T}^0(\delta), \mathcal{T}^1(\delta), \ldots, \mathcal{T}^m(\delta)$ constructed inductively from the π-pruned sequence $\delta' = \langle u'_1, \ldots, u'_m \rangle$ obtained from δ, as follows:

- $\mathcal{T}^0(\delta)$ is empty;
- $\mathcal{T}^1(\delta)$ consists solely of the root $([\{\}, \{s\}], P, 1)$;

[6]Multi-sets are needed to prove Lemma 1, as this relies upon Corollary 6.1 of [Dung *et al.*, 2006] that is proven for multi-sets.

- let $\mathcal{T}^i(\delta)$ be the i-th tree and $u'_{i+1} = \langle \_, \_, t, C, id\rangle$, for $0<i<m$; then $\mathcal{T}^{i+1}(\delta)$ is $\mathcal{T}^i(\delta)$ with additional nodes $([{}_1S_{md},{}_1S_{umd}], L, id), \ldots, ([{}_kS_{md},{}_kS_{umd}], L, id)$ children of the existing nodes $([{}_1S^*_{md},{}_1S^*_{umd}], L^*, U^*_1), \ldots, ([{}_kS^*_{md},{}_kS^*_{umd}], L^*, U^*_k)$ resp., where $k > 0$ and $U^*_1, \ldots, U^*_k$ are such that
  - if $C{=}rl(\rho)$ and there is $u'_j{=}\langle \_, \_, t, rl(\rho'), \_\rangle$ ($0 < j \leq i$) in δ' with $\rho' \neq \rho$ but $Head(\rho){=}Head(\rho')$, then let $u'_{j_1} = \langle \_, \_, t, rl(\rho'_1), jid_1\rangle$, …, $u'_{j_k} = \langle \_, \_, t, rl(\rho'_k), jid_k\rangle$ be all utterances in δ' with target t, $0 < j_l \leq i$, and $\rho'_l \neq \rho$ but $Head(\rho'_l) = Head(\rho)$, for $l = 1, \ldots, k$; then $U^*_1, \ldots, U^*_k$ are the u-ids of the parents in $\mathcal{T}^i(\delta)$ of the nodes $([\_, \_], \_, jid_l)$, $l = 1, \ldots, k$;
  - otherwise, $U^*_1, \ldots, U^*_k$ are the u-ids of the leaf nodes in $\mathcal{T}^i(\delta)$ which are descendants of nodes of the form $([\_, \_], \_, t)$;

  and $([{}_jS_{md}, {}_jS_{umd}], L, id)$ ($1 \leq j \leq k$) are such that
  - if $C = asm(a)$ then
    ${}_jS_{md} ={}_j S^*_{md} \cup \{a\}$,
    ${}_jS_{umd} ={}_j S^*_{umd} \setminus \{a\}$,
    $L = L^*$;
  - if $C = rl(\rho)$ then
    ${}_jS_{md} ={}_j S^*_{md}$,
    ${}_jS_{umd} = ({}_jS^*_{umd} \setminus \{Head(\rho)\}) \cup \{Body(\rho)\}$,
    $L = L^*$;
  - if $C = ctr(a, c)$ then
    ${}_jS_{md} = \{\}$,
    ${}_jS_{umd} = \{c\}$,
    $L = P$ if $L^* = O$, $L = O$ if $L^* = P$.

As an example, figure 1 (Left)[7] gives the dialectical tree drawn from the dialogue in example 3.1.

Definition 9. Given a dialectical tree $\mathcal{T}(\delta)$,

- the *defence set* $\mathcal{D}ef(\mathcal{T}(\delta))$ is the union of all marked sentences from proponent nodes;
- the *culprits* $\mathcal{C}ul(\mathcal{T}(\delta))$ are given by the set of all marked sentences s in opponent nodes N such that the child of N in $\mathcal{T}(\delta)$ is $([\{\}, \{c\}], P, \_)$, and $\mathcal{C}_\delta(s) = c$ in the ABA framework $\langle \mathcal{L}, \mathcal{R}_\delta, \mathcal{A}_\delta, \mathcal{C}_\delta\rangle$ drawn from δ.

Arguments (in the ABA framework drawn from a dialogue) can be drawn from a dialectical tree, as follows:

Definition 10. A *potential argument drawn from a dialectical tree* $\mathcal{T}(\delta)$ is $S{\vdash}_{S'}c$ such that

- $N = ([S, S'], L, \_)$ is a node of $\mathcal{T}(\delta)$,
- $N' = ([\{\}, \{c\}], L, \_)$ is the closest ancestor of N (possible $N' = N$) such that the parent of N' has a label L' different from L, and
- every child of N has a label different from L

An *actual argument drawn from* $\mathcal{T}(\delta)$ is a potential argument $S{\vdash}_{\{\}}c$ drawn from $\mathcal{T}(\delta)$.

[7]Here $([S, S'], L, U)$ with $S = \{e_1, \ldots, e_k\}$ and $S' = \{e'_1, \ldots, e'_h\}$ is represented as $e^m_1, \ldots, e^m_k, e'_1, \ldots, e'_h$:$\mathbf{L}[U]$.

Example 4.1. The potential arguments drawn from the dialogue δ in example 3.1 are: $\{a\}{\vdash}_{\{\}}k$, $\{b\}{\vdash}_{\{\}}q$, $\{c\}{\vdash}_{\{\}}c$ and $\{\}{\vdash}_{\{r\}}r$. The first three are actual arguments. The latter can be seen as a start on the construction of an argument, where only the claim is known. If we consider the same dialogue δ until the utterance with identifier 10, and then:

a_1	a_2
$\langle a_1, a_2, 5, rl(q \leftarrow h), 11\rangle$	$\langle a_2, a_1, 11, rl(h \leftarrow d, g), 12\rangle$
$\langle a_1, a_2, 12, asm(d), 13\rangle$	$\langle a_2, a_1, 13, ctr(d, r), 14\rangle$

we obtain the dialectical tree given in figure 1 (Middle). Here, $\{d\}{\vdash}_{\{g\}}q$ is a new potential argument that is not an actual argument. Note that nodes may hold empty S_{md} and S_{umd} components, as in figure 1.

Some potential arguments correspond to ABA arguments:

Proposition 2. Let $\langle \mathcal{L}, \mathcal{R}_\delta, \mathcal{A}_\delta, \mathcal{C}_\delta\rangle$ be the ABA framework drawn from an ABA dialogue δ and $S{\vdash}_{S'}c$ a potential argument drawn from $\mathcal{T}(\delta)$. If $S' \subseteq \mathcal{A}_\delta$ then $S \cup S' \vdash_{\{\}} c$ is an ABA argument w.r.t. $\langle \mathcal{L}, \mathcal{R}_\delta, \mathcal{A}_\delta, \mathcal{C}_\delta\rangle$.

Thus, actual arguments are necessarily ABA arguments.

We consider now restricted forms of dialectical trees, that we then use below to refine our notion of legal-move function (and to prove Lemma 1).

Definition 11. A dialectical tree $\mathcal{T}(\delta)$ is

- *patient* iff for all nodes $N = ([\_, S_{umd}], L, \_)$ in $\mathcal{T}(\delta)$, N has a child $N' = ([\_, \_], L', \_)$, $L' \neq L$, iff $S_{umd} = \{\}$;
- *focused* iff

  1. each proponent node $([S_{md}, \_], P, \_)$ has either no child, or a single proponent child, or one opponent child per assumption in S_{md}, and

  2. each opponent node has either no child, or any number of opponent children, or one single proponent child.

Basically, arguments in a patient tree are fully expanded (cf. actual) before being attacked. The left tree in figure 1 is patient, whereas the middle tree is not (since the opponent node 13 has a proponent child even though its unmarked component is $\{g\}$). In focused trees, no alternative ways to support or defend claims are considered simultaneously. Both left and middle tree in figure 1 are focused. The following example gives a non-focused tree.

Example 4.2. Let us consider the dialogue in example 3.1 until the utterance with identifier 10, and then:

a_1	a_2
$\langle a_1, a_2, 1, rl(k \leftarrow d), 11\rangle$	$\langle a_2, a_1, 11, asm(d), 12\rangle$
$\langle a_1, a_2, 12, ctr(d, r), 13\rangle$	

The resulting dialectical tree, in figure 1 (Right), is not a focused dialectical tree. Indeed, its root has two proponent children. Accordingly, the rightmost, newly added branch considers an alternative defence (given by $\{d\}$) to the defence (given by $\{a, d\}$) considered in the left-most branch.

The restricted form of legal-move function we consider is guaranteed to generate patient, focused trees, as follows:

Definition 12. A (flat) legal-move function $\lambda : \mathcal{D} \mapsto \mathcal{U}$ is a *patient and focused legal-move function* iff for every dialogue

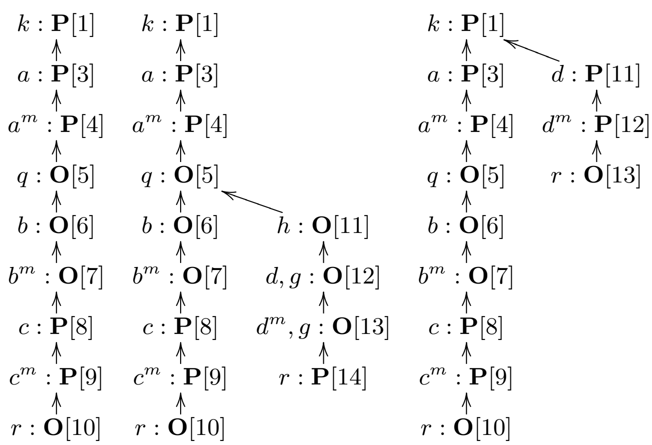

Figure 1: Dialectical trees drawn from the dialogues in examples 3.1 (Left), 4.1 (Middle) and 4.2 (Right).

$\delta \in \mathcal{D}$ such that $\mathcal{T}(\delta)$ is patient and focused, $\mathcal{T}(\delta')$ is still patient and focused, where δ' is the dialogue δ extended with the utterance $\lambda(\delta)$ (represented as $\delta \circ \lambda(\delta)$: if $\delta = \langle u_1, \ldots, u_n \rangle$, then $\delta' = \delta \circ \lambda(\delta) = \langle u_1, \ldots, u_n, \lambda(\delta) \rangle$).

Thus, when an agent decides what to utter, it needs to keep the current dialectical tree into account and make sure that its new utterance keeps the tree patient and focused. As a result, the dialectical tree can be seen as a *commitment store*.

Our dialogue model so far does not force agents to say *legitimate utterances possible* given the dialogue so far, where, for example, in the ABA dialogue δ^*

a_1	a_2
$\langle a_1, a_2, 0, clm(a), 1 \rangle$	$\langle a_2, a_1, 1, asm(a), 2 \rangle$
$\langle a_1, a_2, 2, ctr(a, p), 3 \rangle$	$\langle a_2, a_1, 3, rl(p \leftarrow c), 4 \rangle$
$\langle a_1, a_2, 4, asm(c), 5 \rangle$	$\langle a_2, a_1, 5, ctr(c, q), 6 \rangle$
$\langle a_1, a_2, 6, rl(q \leftarrow), 7 \rangle$	$\langle a_2, a_1, 3, rl(p \leftarrow q), 8 \rangle$

$u^* = \langle a_1, a_2, 8, rl(q \leftarrow), 9 \rangle$ is a possible legitimate utterance, since $q \leftarrow \in \mathcal{R}_\delta$ due to utterance 7.

Definition 13. The *exhaustive outcome function* ω_{ex} is such that, given $\delta \in \mathcal{D}$ and $\langle \mathcal{L}, \mathcal{R}_\delta, \mathcal{A}_\delta, \mathcal{C}_\delta \rangle$ the framework drawn from δ, $\omega_{ex}(\delta) = true$ iff $\omega_{ABA}(\delta) = true$ and $\nexists u'$ with content either $asm(a)$, for $a \in \mathcal{A}_\delta$, or $rl(\rho)$, for $\rho \in \mathcal{R}_\delta$, or $ctr(a, c)$, for $c = \mathcal{C}_\delta(a)$, such that $\delta \circ u'$ is an ABA dialogue.

Thus, $\omega_{ex}(\delta^*)$=$false$ and $\omega_{ex}(\delta^* \circ u^*)$=$true$. Note that exhaustiveness does not force complete dialogues or that agents contribute to dialogues with relevant information they hold.

We refer to ABA dialogues generated by a patient, focused legal-move function and for which ω_{ex} is true as *exhaustive*.

5 Formal results

In this section we link dialogues and the admissible argumentation semantics. First we refine the outcome function so that if a dialogue has a true outcome then the (fictional) proponent has the last word in the dialogue, namely all leaves in the dialectical tree are proponent nodes or "dead-end" opponent nodes (not corresponding to any actual argument). Formally:

Definition 14. The *last word outcome function* ω_{lw} is such that, given $\delta \in \mathcal{D}$, $\omega_{lw}(\delta) = true$ iff $\omega_{ex}(\delta) = true$, there is no non-empty proponent leaf and no empty opponent leaf in $\mathcal{T}(\delta)$, where a node $([S_1, S_2], \_, \_)$ is *empty* iff $S_1 \cup S_2 = \{\}$.

We refer to exhaustive dialogues for which ω_{lw} is true as *positive*. The dialogue in example 3.1 is positive. Positive dialogues give dialectical trees corresponding to *concrete dispute tree* with the same defence set and culprits. Formally:

Lemma 1. Given a dialogue $\delta = D^{a_j}_{a_i}(s)$ such that $\omega_{lw}(\delta) = true$, there is a concrete dispute tree $\mathcal{T}^c$ for s, such that $\mathcal{D}ef(\mathcal{T}(\delta)) = \mathcal{D}ef(\mathcal{T}^c)$ and $\mathcal{C}ul(\mathcal{T}(\delta)) = \mathcal{C}ul(\mathcal{T}^c)$.

Proof. (Sketch) 1) Given that $n \geq 1$ is the number of nodes in $\mathcal{T}(\delta)$, we assign a unique number $1 \leq i \leq n$ to each node of $\mathcal{T}(\delta)$ so that if $i < j$ then N_i is an ancestor of N_j (the root is $N_1 = ([\{\}, \{s\}], P, 1)$). We then construct the sequence $\mathcal{T}'_0, \mathcal{T}'_1, \ldots, \mathcal{T}'_n$ such that $\mathcal{T}'_0$ is empty, $\mathcal{T}'_1$ consists solely of the root $N'_1 = (P, \{s\})$, and, for $0 < i < n$, $\mathcal{T}'_{i+1}$ is $\mathcal{T}'_i$ with an extra node N'_{i+1} child of the node N'_k corresponding to N_k in $\mathcal{T}(\delta)$ that is the parent of N_{i+1}. If N_{i+1}=$([S^{i+1}_{md}, S^{i+1}_{umd}], L^{i+1}, \_)$, then N'_k=(L', S') with $L' = L^{i+1}$ and S'=$S^{i+1}_{md} \cup S^{i+1}_{umd}$. 2) We remove all duplicate nodes in $\mathcal{T}'_n$, getting $\mathcal{T}''$. 3) We transform $\mathcal{T}''$ so that each proponent node has at most one opponent node as a child, getting $\mathcal{T}^c$. 4) We prove that $\mathcal{T}^c$ is a concrete dispute tree. □

As in the case of concrete dispute trees, the defence set of a dialectical tree may not be admissible, as it may attack itself. We refine the notion of outcome function by enforcing that this set does not attack itself, as follows:

Definition 15. The *admissible outcome function* ω_{ADM} is such that, given $\delta \in \mathcal{D}$, $\omega_{ADM}(\delta) = true$ iff $\omega_{lw}(\delta) = true$ and $\mathcal{D}ef(\mathcal{T}(\delta)) \cap \mathcal{C}ul(\mathcal{T}(\delta)) = \{\}$. If δ is positive and $\omega_{ADM}(\delta) = true$, we say that δ is *successful*.

Theorem 1. Given a *successful* dialogue $D^{a_j}_{a_i}(s) = \delta$, then there exists an argument $S \vdash s$ that belongs to an *admissible* extension supported by $\mathcal{D}ef(\mathcal{T}(\delta))$ w.r.t. the ABA framework drawn from δ.

Proof. If $\omega_{ADM}(\delta) = true$, by Lemma 1 there exists a concrete dispute tree $\mathcal{T}^c$ such that $\mathcal{D}ef(\mathcal{T}(\delta)) = \mathcal{D}ef(\mathcal{T}^c)$ and $\mathcal{C}ul(\mathcal{T}(\delta)) = \mathcal{C}ul(\mathcal{T}^c)$. By Corollary 6.1 of [Dung *et al.*, 2006] (see section 2), the theorem holds. □

6 Illustration

Two agents, Jenny (**J**) and Amy (**A**), are planning a film night together and want to agree on a movie to watch. The agreement can be reached through a dialogue, as follows:

J: Let's see if *Terminator* is a good movie to watch.
A: OK.
J: I want a movie that is fun and has a good screening time.
A: OK.
J: To me, a movie is fun if it is an action movie. **A:** OK.
J: And, *Terminator* is an action movie. **A:** OK.
J: I also believe *Terminator* starts at the right time.
A: Are you sure it is not going to be too late? **J:** Why?
A: I don't know. I am just afraid so.
J: It won't be too late if it finishes by 10 o'clock.
A: I see. Indeed *Terminator* finishes by 10 o'clock.

J: OK.
A: OK.

Here Jenny succeeds in persuading Amy to watch the movie she proposes. Amy had the opportunity to disagree and challenge Jenny, but Jenny managed to produce a compelling argument. This natural language dialogue gives an informal reading of the dialogue in our framework.[8]

$\langle J, A, 0, clm(wM(t)), 1\rangle$	$\langle A, J, 0, \pi, 2\rangle$
$\langle J, A, 1, rl(wM(t) \leftarrow f(t), sT(t)), 3\rangle$	$\langle A, J, 0, \pi, 4\rangle$
$\langle J, A, 3, rl(f(t) \leftarrow aM(t)), 5\rangle$	$\langle A, J, 0, \pi, 6\rangle$
$\langle J, A, 5, rl(aM(t)), 7\rangle$	$\langle A, J, 0, \pi, 8\rangle$
$\langle J, A, 3, asm(sT(t)), 9\rangle$	
$\langle A, J, 9, ctr(sT(t), late(t)), 10\rangle$	$\langle J, A, 0, \pi, 11\rangle$
$\langle A, J, 10, asm(late(t)), 12\rangle$	
$\langle J, A, 12, ctr(late(t), fT(t)), 13\rangle$	
$\langle A, J, 13, rl(fT(t)), 14\rangle$	$\langle J, A, 0, \pi, 15\rangle$
	$\langle A, J, 0, \pi, 16\rangle$

This dialogue is successful and, by theorem 1, $wM(t)$ belongs to an admissible extension of the drawn ABA framework. This corresponds to Jenny having persuaded Amy, in that no objections have been raised that could not be addressed, and Jenny's view point is non-contradictory.

7 Related Work

[McBurney and Parsons, 2009] give an overview of dialogue games for argumentation. Our work can be seen as providing a novel dialogue game for ABA. [Prakken, 2006] reviews dialogue systems for persuasion. Section 6 illustrates how our model can be used for persuasion. However, it can also be used for other kinds of dialogues, e.g. information-seeking.

[Parsons *et al.*, 2007] examine three notions of relevance in dialogues where utterances are arguments and attacks. Our utterances are at at a finer granularity level, as they correspond to rules etc. Thus, there is no direct mapping between our work and their relevance.

[Black and Hunter, 2009] present a formal system for inquiry dialogues based on DeLP as the underlying argumentation framework. Our work differs from theirs as (1) it does not focus on inquiry dialogues; (2) it uses ABA; (3) it does not force full disclosure of all relevant knowledge as they do.

[Prakken, 2005] defines a formal system for persuasion. The main differences with our work are: (1) proponent and opponent roles are pre-assigned to agents before the dialogue whereas in our work agents can play both roles within the same dialogue; (2) he considers the grounded semantics, whereas we use admissibility; (3) his set of utterances refer to arguments and attacks, as in the case of [Parsons *et al.*, 2007]; (4) he forces the support of arguments to be minimal, whereas we do not, in the spirit of [Dung *et al.*, 2010]; (5) he does not impose a strict interleaving, whereas we do, similarly to [Black and Hunter, 2009]. (6) he does not allow agents to jointly construct arguments whereas we do.

[8] wM stands for watchMovie, t for Terminator, sT for screenTime, f for fun, aM for actionMovie, fT for finishbyTen.

8 Conclusion

We presented a formal dialogue model for assumption-based argumentation (ABA). The model is not tailored to any dialogue type. We chose to illustrate the model with a persuasion dialogue. We proved that our model is sound by connecting it with the admissibility argumentation semantics for ABA. Thus, our dialogues can be seen as a distributed mechanism for computing admissible extensions (supporting claims).

We have assumed that agents exchange their views in ABA-format, namely ABA serves as a standard for the exchange of information between agents. However, agents may adopt an internal representation different from ABA.

Future work includes further investigating properties of model, including soundness w.r.t. other argumentation semantics, as well as completeness. We are also interested in extending the model to allow for a more flexible turn-making than strict interleaving, to support conversations amongst more than two agents, and to refine the model as to prevent undesirable strategic behavior of agents, e.g. along the lines of [Rahwan *et al.*, 2009]. Finally, we plan to develop applications of our model for the various kinds of dialogue types, e.g. inquiry, information-seeking and negotiation.

References

[Black and Hunter, 2009] E. Black and A. Hunter. An in quiry dialogue system. *JAAMAS*, 19:173–209, 2009.

[Bondarenko *et al.*, 1997] A. Bondarenko, P. M. Dung, R. Kowalski, and F. Toni. An abstract, argumentation-theoretic approach to default reasoning. *Art. Int.*, 93(1-2):63–101, 1997.

[Dung *et al.*, 2006] P.M. Dung, R.A. Kowalski, and F. Toni. Dialectic proof procedures for assumption-based, admissible argume ntation. *Artificial Intelligence*, 170:114–159, 2006.

[Dung *et al.*, 2009] P. M. Dung, R. A. Kowalski, and F. Toni. Assumption-based argumentation. In *Argumentation in Artificial Intelligence*, pages 25–44. Springer, 2009.

[Dung *et al.*, 2010] P. M. Dung, F. Toni, and P. Mancarella. Some design guidelines for practical argumentation systems. In *Proc. COMMA*, pages 183–194, 2010.

[McBurney and Parsons, 2009] P. McBurney and S. Parsons. Dialogue games for agent argumentation. In *Argumentation in Artificial Intelligence*, pages 261–280. 2009.

[Parsons *et al.*, 2007] S. Parsons, P. McBurney, E. Sklar, and M. Wooldridge. On the relevance of utterances in formal inter-agent dialogues. In *Proc. AAMAS*, pages 47–62, 2007.

[Prakken, 2005] H. Prakken. Coherence and flexibility in dialogue games for argumentation. *JLC*, 15:1009–1040, 2005.

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

[Rahwan *et al.*, 2009] I. Rahwan, K. Larson, and F. A. Tohmé. A characterisation of strategy-proofness for grounded argumentation semantics. In *Proc. IJCAI*, pages 251–256, 2009.