A Sound and Complete Dialogue System for Handling Misunderstandings

Andreas Xydis[1,†], Christopher Hampson[1], Sanjay Modgil[1] and Elizabeth Black[1]

[1]*Department of Informatics, King's College London, London, United Kingdom*

Abstract

A common assumption for argumentation dialogues is that any argument exchanged is complete, i.e. its premises entail its claim. But in real world dialogues, agents commonly exchange enthymemes — arguments with incomplete logical structure. This paper introduces a dialogue system that handles misunderstandings that may occur when there is a mismatch between what a participant excludes from an argument and what her interlocutor assumes to be the missing information. We also introduce a mechanism for determining the acceptability status of the dialogue moves so that, under certain conditions, the status of moves made during a dialogue conforming to our system, corresponds with the status of arguments in the Dung argument framework instantiated by the contents of the moves made at that stage in the dialogue.

Keywords

Enthymemes, locutions, dialogue, framework, ASPIC$^+$, argumentation

1. Introduction

In approaches to structured argumentation, arguments typically consist of a conclusion deductively and/or defeasibly inferred from some premises [1]. However, in practice, human agents typically assert 'incomplete' arguments known as enthymemes [2]. Misunderstandings may arise due to interlocutors' incorrect assumptions. Consider for example the dialogue below, which is annotated with the relevant locutions from the dialogue system proposed in this paper.

Example 1. **1. Alice:** *I found a job so I can eat at a restaurant today. (assert $b; b \Rightarrow a; a$)*
2. Bob: *You can't afford to eat out, you owe money. (assert $c; c \Rightarrow \neg a; \neg a$)*
3. Alice: *But I made a deal with my creditors. (assert q)*
4. Bob: *And so? What do you mean? (what do you intend)*
5. Alice: *So I don't need to pay the bills today. (intend $q; q \Rightarrow \neg e; \neg e$)*
6. Bob: *Why is that relevant? (what did you assume by $c; c \Rightarrow \neg a; \neg a$)*
7. Alice: *I thought that you thought that I can't afford to eat at a restaurant today because I owe money since I have bills to pay today. (assumed $e; e \rightarrow c; c; c \Rightarrow \neg a; \neg a$)*
8. Bob: *No! I meant that you can't afford to eat at a restaurant today because you owe money since you need to pay Kate back today. (meant $p; p \rightarrow c; c; c \Rightarrow \neg a; \neg a$)*

SAFA'22: Fourth International Workshop on Systems and Algorithms for Formal Argumentation 2022, September 13, 2022, Cardiff, Wales, United Kingdom

† Corresponding author E-mail: andreas.xydis@kcl.ac.uk

© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)
CEUR-WS.org/Vol-3236/paper2.pdf

Alice first asserts an argument (1), to which Bob responds with a counterargument (2). Alice then moves an enthymeme (3) for an argument that she believes counters Bob's argument. Note, the enthymeme Alice presents does not explicitly contradict anything that Bob has said, and so Bob asks for clarification (4) on what he is meant to infer from this enthymeme, which Alice provides (5)[1]. However, Alice's clarification still does not explicitly contradict anything Bob has said. Since Bob does not understand why Alice's enthymeme is relevant to what he said, he asks Alice to explain what she assumed (6). Alice explains the assumption she had made (7), which Bob then corrects (8). Notice that Bob's assertion in (2) is an enthymeme of his intended argument in (8).

This simple example illustrates the need for a dialogue system that allows agents to ask for and provide clarifications when enthymemes are involved (4 and 5, respectively). It also shows that agents need to be able to ask what another agent has assumed was intended by an enthymeme (6), to answer such a question (7), and to correct any erroneous assumptions (8). Additionally, observe that before Bob resolves the misunderstanding regarding what he intended and what Alice assumed (6-8), although it appears that Alice wins the dialogue (since she moves her enthymeme in 3 against Bob's argument in 2), there is no formal defeat on Bob's argument and so Bob's argument (2) is determined acceptable in the argument framework AF constructed by the contents of the enthymemes revealed by the agents. In other words, a mismatch can exist between the pragmatic and the logical conclusions implied by a dialogue in which enthymemes are used.

Most works dealing with enthymemes focus on how an enthymeme can be constructed from the intended argument and how the intended argument can be reconstructed from a received enthymeme, based on assumptions that the sender and the receiver make about their shared knowledge and context, e.g. [4, 5, 6]. Notice that [7] allows for both backward and forward extending of enthymemes, as does the dialogue system in [8], which additionally enables resolution of misunderstandings that arise due to use of enthymemes. However, unlike our work, these works do not consider how the outcome of the dialogue relates to the AF that is instantiated based on contents of the enthymemes moved during a dialogue, meaning that there is no guarantee that the dialogue outcome respects the underlying argumentation theory. Notable exceptions are [9] and [10], but these works only address backward and forward extension of enthymemes, respectively, whereas our paper allows for the full extension of any kind of enthymeme to the intended argument it was instantiated from as well as for the resolution of misunderstandings that may occur because of enthymemes.

Our primary contribution is the development of a dialogue system in which agents can move any kind of enthymeme and seek clarification to elicit the intended arguments of these enthymemes as well as resolve any misunderstanding that may occur between the participants regarding what an agent assumes and what its counterparts intends, such that under certain conditions, the dialogical status of the moves made during the dialogue—determined by what we call the *dialogue framework*—corresponds to the acceptability of the arguments in the Dung AF instantiated from the declarative contents of the moves made at that stage in the dialogue. This correspondence—not considered previously for dialogue systems that support

[1]This example highlights requirements for dialogical 'reification' of relations between locutions, as proposed in [3]; requirements that arise due to the use of enthymemes

resolution of misunderstandings in dialogues—is significant since it demonstrates that the dialogue system we propose respects the logic and semantics of the underlying argumentation theory. Essentially, this correspondence means that there is no disadvantage to the use of enthymemes in dialogues (even when misunderstandings occur); a common real-world feature of dialogues that supports efficient inter-agent communication. If we are to enable effective human-computer interaction and provide normative support for human-human dialogue, we need to account for the ubiquitous use of enthymemes in real-world dialogues and make sure that agents can still reach the “correct" conclusions based on the knowledge they have shared.

2. Preliminaries

The $ASPIC^+$ framework [11] abstracts from the particularities of the underlying language, the nature of conflict, the defeasible and strict inference rules (which can encode a deductive logic of one's choosing) that are used to chain inferences from premises to an argument's conclusion. $ASPIC^+$ arguments are evaluated in a Dung AF [12], and $ASPIC^+$ provides guidelines for ensuring that the outcome of evaluation yields rational outcomes.

An $ASPIC^+$ *argumentation theory* AT is a tuple $\langle AS, K \rangle$ consisting of an *argumentation system* AS and a *knowledge base* K. AS is a tuple $\langle L, (\bar{\cdot}), R, nom \rangle$ where L is a logical language and $(\bar{\cdot}) : L \to (2^L - \{\emptyset\})$ is a function that generalises the notion of negation, so as to declare that two formulae are in conflict (e.g., $\overline{\text{married}} = \{\text{single}, \text{unmarried}\}$). Additionally, $R = R_s \cup R_{def}$ is a set of strict (R_s) and defeasible (R_{def}) inference rules and $nom : R^*_{def} \to L$ (where R^*_{def} is the class of all defeasible rules) is a naming function which assigns a name (or nominal) to each defeasible rule (so that rules can be referenced in the object language). Lastly, $R = R_s \cup R_{def}$ is the (disjoint) union of a set of *strict rules* R_s of the form $p_1, \ldots, p_n \to p$, and a set of *defeasible rules* R_{def} of the form $p_1, \ldots, p_n \Rightarrow p$, for $p, p_1, \ldots, p_n \in L$. For each rule $r \in R$ let $\mathsf{antecedents}(r) = \{p_1, \ldots, p_n\}$ denote the set of antecedents of r and $\mathsf{consequent}(r) = p$ denote the consequent of r. Finally, a knowledge base $K \subseteq L$ is a set of premises.

In our representation of an *argument*, a strict/defeasible inference rule is incorporated as a node, intermediating between the parent node (the rule's conclusion) and the child node(s) (the rule's antecedent(s)). This contrasts with the standard $ASPIC^+$ notion of an argument in which the rules are represented by undirected edges (where solid lines represent strict rules and dotted lines represent defeasible rules) linking the conclusion to the rule's antecedents (see Figures 1.(a) and 1.(b) for an example of an $ASPIC^+$ argument and an argument as depicted in this paper, respectively).

Definition 1. *Given an argumentation system* $AS = \langle L, (\bar{\cdot}), R, nom \rangle$ *and an argumentation theory* $AT = \langle AS, K \rangle$, *an* ***argument*** *is a labelled (downward directed) tree* $A = \langle \mathsf{Nodes}(A), \mathsf{Edges}(A), \mathsf{lab}_A \rangle$ *such that:*

1. $\mathsf{lab}_A : \mathsf{Nodes}(A) \to L \cup R$ *is a node labelling;*
2. $\mathsf{Edges}(A) \subseteq \mathsf{Nodes}(A) \times \mathsf{Nodes}(A)$ *such that if* $(n_i, n_j) \in \mathsf{Edges}(A)$, *then either:*
   (a) $\mathsf{lab}_A(n_i) \in L, \mathsf{lab}_A(n_j) \in R$ *and* $\mathsf{consequent}(\mathsf{lab}_A(n_j)) = \mathsf{lab}_A(n_i)$, *or*
   (b) $\mathsf{lab}_A(n_i) \in R$ *and* $\mathsf{antecedents}(\mathsf{lab}_A(n_i)) = \{\mathsf{lab}_A(n_j) \mid (n_i, n_j) \in \mathsf{Edges}(A)\}$,

and if $(n_i, n_j), (n_i, n_k) \in \mathsf{Edges}(A)$ *and* $n_j \neq n_k$ *then* $\mathsf{lab}_A(n_j) \neq \mathsf{lab}_A(n_k)$*;*

3. *for every node* $n \in \mathsf{Leaves}(A)$*, we have* $\mathsf{lab}_A(n) \in K$*;*
4. $|\mathsf{Roots}(A)| = 1$*;*
5. $\mathsf{lab}_A(\mathsf{Conc}(A)) \in L$*, where* $\mathsf{Conc}(A)$ *is the unique element in* $\mathsf{Roots}(A)$*;*

where $\mathsf{Leaves}(A) = \{n_i \in \mathsf{Nodes}(A) \mid \nexists(n_i, n_j) \in \mathsf{Edges}(A)\}$, $\mathsf{Roots}(A) = \{n_i \in \mathsf{Nodes}(A) \mid \nexists(n_j, n_i) \in \mathsf{Edges}(A)\}$*. Let* A_{AT} *denote the set of arguments instantiated by* AT*,* A^* *the class of all arguments, and* $\mathsf{Rules}(A) = \{n \in \mathsf{Nodes}(A) \mid \mathsf{lab}_A(n) \in R\}$*, for all* $A \in A^*$*.*

The *attack* and *defeat* relations over arguments are defined as for $ASPIC^+$ arguments. Argument X attacks argument Y if X's claim conflicts with an ordinary premise or the consequent or name of a defeasible rule in Y, and may succeed as a *defeat* if the targeted sub-argument Y' of Y ($Y' \in \mathsf{Sub}(X)$) is not strictly preferred to X (for details see [11]). An *a*rgument framework AF *i*nstantiated by an argumentation theory AT is a tuple $\langle A_{AT}, \mathit{Dfs}\rangle$ where $\mathit{Dfs} \subseteq A_{AT} \times A_{AT}$ is the $ASPIC^+$ *defeat relation.* According to a *complete labelling* function L on AF (as defined in [13]) $\forall A \in A_{AT}$: A is IN iff every B that defeats A is OUT ($(B, A) \in \mathit{Dfs} \rightarrow B$ is OUT); A is OUT iff $\exists(B, A) \in \mathit{Dfs}$, B is IN; UNDEC iff A is neither IN nor OUT. $\mathsf{in}(\mathsf{L})$, $\mathsf{out}(\mathsf{L})$ and $\mathsf{undec}(\mathsf{L})$ respectively denote the set of arguments that are IN, OUT and UNDEC. A *preferred labelling* L_{pr} is a complete labelling where $\mathsf{in}(\mathsf{L}_{pr})$ is maximal, the *grounded labelling* L_{gr} is a complete labelling where $\mathsf{in}(\mathsf{L}_{gr})$ is minimal and a *stable labelling* L_{st} is a complete labelling where $\mathsf{undec}(\mathsf{L}_{pr})$ is empty.

Enthymemes are incomplete arguments. Contrary to other approaches that handle enthymemes [14, 15], we allow omission of an argument's claim, as well as its premises, and so may obtain a disjointed graph (as the claim is the root of the tree that is the intended argument). Hence we represent enthymemes as a forest of trees (see Figure 1.(c) for an example). Since an *enthymeme* E is constructed from an intended argument A, one may choose to remove the conclusion of an inference rule while retaining the inference rule, or remove a sub-argument whose conclusion is the antecedent of a strict/defeasible rule. So, an enthymeme E of an argument A is a forest of trees whose nodes and edges are a subset of the nodes and edges of A, and the label of each node in E is the same label of the corresponding node in A (for example see X' in Figure 1.(b) and Figure 1.(c)).

Definition 2. *Let* $AS = \langle L, \bar{(\cdot)}, R, nom\rangle$*. An **enthymeme** E is a forest of labelled (downward directed) trees* $E = \langle \mathsf{Nodes}(E), \mathsf{Edges}(E), \mathsf{lab}_E\rangle$ *such that* $\mathsf{Nodes}(E) \neq \emptyset$ *and conditions 1 and 2 from Definition 1 hold.*

We say that E *is* an enthymeme of an argument A *(written* $E \leq A$*) if* $\mathsf{Nodes}(E) \subseteq \mathsf{Nodes}(A)$, $\mathsf{Edges}(E) \subseteq \mathsf{Edges}(A)$ *and* $\mathsf{lab}_E(n) = \mathsf{lab}_A(n)$*, for all* $n \in \mathsf{Nodes}(E)$*. Then* $\mathsf{Roots}(E) = \{n_i \in \mathsf{Nodes}(E) \mid \nexists(n_j, n_i) \in \mathsf{Edges}(E)\}$ *and* $\mathsf{Leaves}(E) = \{n_i \in \mathsf{Nodes}(E) \mid \nexists(n_i, n_j) \in \mathsf{Edges}(E)\}$.

Note that an argument A is an enthymeme of itself, where $\mathsf{Roots}(A) = \{\mathsf{Conc}(A)\}$ and $\{\mathsf{lab}(n) : n \in \mathsf{Leaves}(A)\} \subseteq K$. Thus, we also write that $E' \leq E$ (where E' and E are enthymemes) to denote that E' is an enthymeme of E (or else E **extends** E', e.g. $B' \leq B$ and B extends B' in Figure 1.(b)). The class of all enthymemes is denoted E^* and so includes the class of all arguments ($A^* \subseteq E^*$).

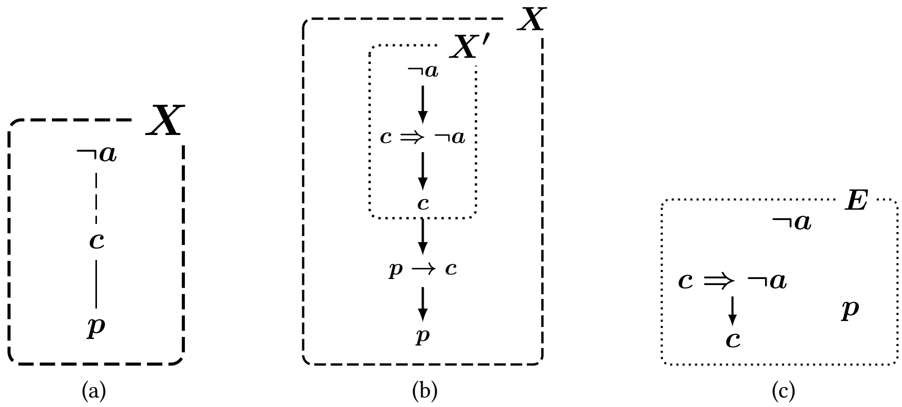

Figure 1: (a) Bob's argument X (enclosed by a dashed line) in line 8 in Example 1 as represented in $ASPIC^{+}$. $\neg a$ is the claim of X and p is a premise that strictly (since connected with a solid line) infers the intermediate claim c, and c defeasibly (since connected with a dashed line) infers $\neg a$; (b) X as represented in this paper, where the inference rules applied ($\Rightarrow$ denotes defeasible inference, $\rightarrow$ strict inference) are explicitly represented and a claim is connected to its premises via the inference rules applied. Notice that X' (enclosed by a dotted line) is the enthymeme moved in line 2 in Example 1. (c) An enthymeme E (of X).

3. Dialogue system

An *enthymeme dialogue* d is a sequence of moves. Each move is a 5-tuple comprising the move's *sender*, *locution*, *content*, *reply* and *target*. The sender is either $Prop$ or Op (the participants of the dialogue). Table 1 shows the locutions permitted in our system.

A move's *content* depends on its locution: if a move's locution is assert or w.d.y.a.b. its content is an enthymeme. If a move's locution is intend, assumed, meant or agree its content is an argument, otherwise the move's content is $\emptyset$. The w.d.y.i., intend, assumed, meant and agree moves explicit responses to a previous move. For any move with one of the aforementioned locutions, its *reply* is a natural number indexing the move to which it replies, otherwise its *reply* is $\emptyset$.

An asserted enthymeme is moved as a defeat against a previously moved enthymeme — the *target* of the assertion — whereas a w.d.y.a.b. move questions the defeat relation between the *content*s of two moves. So, if a move's locution is assert, its *target* is the natural number indexing the move whose content is the enthymeme being defeated, whereas if a move's locution is w.d.y.a.b., its *target* is the pair of natural numbers (i, j) where m_i targets m_j (m_i's enthymeme is moved as a defeat on m_j's enthymeme), otherwise its target is $\emptyset$.

When an agent asserts an enthymeme, reveals its intended argument, corrects or confirms its counterpart's interpretation of an enthymeme it has in mind a complete argument and this is the *intended argument* of the move. Note that since we accommodate nefarious agents, we do not insist that a move's intended argument necessarily extends the move's enthymeme. Similarly, when an agent asserts an enthymeme E so as to defeat some targeted enthymeme E', the agent has in mind an argument A that it believes was intended as the complete argument extending

Locution	Meaning
assert	Assert an enthymeme.
w.d.y.i.	Request the intended argument of an asserted enthymeme by asking *"what do you intend"*.
intend	Provide the intended argument of an asserted enthymeme.
w.d.y.a.b.	Check the other participant's understanding of an enthymeme by asking *"what did you assume by . . . "*.
assumed	Provide their own interpretation of an enthymeme.
meant	Correct the other participant's interpretation of an enthymeme.
agree	Confirm the other participant's interpretation of an enthymeme.
stop	Indicate the participant's intention to stop the dialogue.

Table 1
Table of locutions permitted in our dialogue system and their meanings.

the targeted E'. A is said to be the assert move's *intended target argument.* See Table 2 for an example of an enthymeme dialogue.

Definition 3. *An* ***enthymeme dialogue*** *between two participants Prop and Op is a sequence of* ***moves*** $d = [m_0, m_1, \ldots, m_\ell]$*, where each move* $m_i = \langle \mathtt{s}(m_i), \mathtt{l}(m_i), \mathtt{c}(m_i), \mathtt{re}(m_i), \mathtt{t}(m_i)\rangle$ *is a 5-tuple comprising the following:*

1. Sender: $\mathtt{s}(m_i) \in \{Prop, Op\}$*;*
2. Locution: $\mathtt{l}(m_i) \in \{\mathsf{assert}, \mathsf{w.d.y.i.}, \mathsf{intend}, \mathsf{w.d.y.a.b.}, \mathsf{assumed}, \mathsf{meant}, \mathsf{agree}, \mathsf{stop}\}$*;*
3. Content: *if* $\mathtt{l}(m_i) \in \{\mathsf{assert}, \mathsf{w.d.y.a.b.}\}$ *then* $\mathtt{c}(m_i) \in E^*$*, else if* $\mathtt{l}(m_i) \in \{\mathsf{intend}, \mathsf{assumed}, \mathsf{meant}, \mathsf{agree}\}$ *then* $\mathtt{c}(m_i) \in A^*$*, otherwise* $\mathtt{c}(m_i) = \emptyset$*;*
4. Reply: *if* $\mathtt{l}(m_i) \in \{\mathsf{w.d.y.i.}, \mathsf{intend}, \mathsf{assumed}, \mathsf{meant}, \mathsf{agree}\}$, $\mathtt{re}(m_i) \in \{0, \ldots, (i-1)\}$*, otherwise* $\mathtt{re}(m_i) = \emptyset$*;*
5. Target: *if* $\mathtt{l}(m_i) \in \{\mathsf{assert}\}$*,* $\mathtt{t}(m_i) \in \{0, \ldots, (i-1)\} \cup \{\emptyset\}$*, otherwise if* $\mathtt{l}(m_i) = \mathsf{w.d.y.a.b.}$*,* $\mathtt{t}(m_i) \in (\{0, \ldots, (i-1)\} \times \{0, \ldots, (i-1)\})$*, otherwise* $\mathtt{t}(m_i) = \emptyset$*.*

We denote the class of all moves as M^**. If* $\mathtt{l}(m_i) \in \{\mathsf{assert}, \mathsf{intend}, \mathsf{meant}, \mathsf{agree}\}$*,* $\mathsf{IntArg}(m_i) \in (A^* \cup \{\emptyset\})$ *is the* ***intended argument*** *of* m_i *and if* $\mathtt{l}(m_i) = \mathsf{assert}$*,* $\mathsf{IntTarArg}(m_i) \in (A^* \cup \{\emptyset\})$ *is the* ***intended target argument****, else* $\mathsf{IntArg}(m_i) = \emptyset$ *and* $\mathsf{IntTarArg}(m_i) = \emptyset$*, respectively. We let* $d_i = [m_0, \ldots, m_i]$ *be the enthymeme dialogue whose last move is* m_i *and say that* $d_k = [m_0, \ldots, m_i, \ldots, m_k]$ **expands** d_i *where* $i < k < \ell$*.*

Abusing notation, we identify $\mathtt{re}(m_i)$ with the move $m_{\mathtt{re}(m_i)}$ to which m_i replies and $\mathtt{t}(m_i)$ with the move $m_{\mathtt{t}(m_i)}$ (or the pair of moves (m_j, m_k)) that m_i targets, instead of the indices $\mathtt{re}(m_i)$ and $\mathtt{t}(m_i)$ (or (j, k)) of $m_{\mathtt{re}(m_i)}$ and $m_{\mathtt{t}(m_i)}$ (or (m_j, m_k)), respectively.

We now define a *well-formed enthymeme dialogue* d with two participants, $Prop$ and Op. We assume that the agents share the same underlying argumentation system, ensuring they can understand each other, except that the defeasible rules they are aware of may differ. They may of course have different knowledge (belief) bases. The participants alternate turns and cannot repeat moves. $Prop$ must start by asserting an *argument* (the conclusion of which we call the *topic* of the dialogue). Apart from this first assert move, which has no target, an assert move's enthymeme targets (i.e., is moved as a defeat against) a previously asserted enthymeme.

Note that while m_i targets m_j, this does not necessarily imply that the content E of m_i is a valid defeat on the content E' of m_j, according to the $ASPIC^+$ definition of defeat. Since enthymemes may only partially reveal the content of an argument, one may not be able to validate that a target relation corresponds to an $ASPIC^+$ defeat; hence our definition of well-formed enthymeme dialogues does not enforce such a correspondence. This of course means that participants may be mistaken in the assumptions they have made regarding the other's intended arguments (i.e., E may not legitimately defeat the intended complete argument A that extends E'), or, for strategic purposes, an agent may dishonestly target a move's enthymeme E', even though they know that their enthymeme E does not defeat the intended target argument A. A w.d.y.i. move is used to request the intended argument of a previously asserted enthymeme E, while an intend move replies to a w.d.y.i. move and supplies the complete argument extending E.

Recall that the receiver of an asserted enthymeme E' might not be able to interpret correctly the intended argument A' extending E'. Thus, if m_j targets m_i, the agent Ag who moved m_i might not be able to understand why the content E' of m_j has been moved against the content E of m_i (i.e., Ag is unable to reconstruct an argument A' from E', such that A' represents a valid defeat on Ag's intended argument A extending E). Moreover, the sender Ag$'$ of m_j might be mistaken in her interpretation of the intended target argument A extending E in m_i. So, Ag can check whether Ag$'$ correctly assumes that A extends E, by making a w.d.y.a.b. move m_k (whose content repeats the content E of m_i) and which essentially amounts to questioning the validity of the targeting relationship (m_j, m_i) (i.e., whether it corresponds to a valid defeat); thus, the target of m_k is (m_j, m_i).

A participant of a *well-formed enthymeme dialogue* d replies to a w.d.y.a.b. move m_k only with an assumed move m_h. We also impose that the content of m_h is an argument that extends the content of m_k since by using m_h its sender (Ag$'$) reveals its understanding of the argument extending Ag's enthymeme (E) in m_k. A reply to the assumed move m_h is a move m_g with locution meant or agree, in which Ag reveals the intended argument A extending the enthymeme E in m_i (and which is repeated in m_k). Now, if Ag$'$ (in the move m_h) was mistaken in its assumption as to what was the intended argument, then m_g's locution is meant and its content A must not be the same as the mistaken content B of m_h. Otherwise, if Ag$'$'s interpretation was correct (i.e., Ag$'$ moved A in m_h), then m_g's locution is agree and its content is A; that is, the argument in m_h and m_g is the intended argument A that extends the enthymeme E in Ag's earlier move m_i.

Finally, d is terminated when two stop moves are made consecutively (indicating that neither agent wishes to continue the dialogue). For an example of a well-formed enthymeme dialogue see Table 2.

Definition 4. *Let* $AS_{\mathsf{Ag}} = \langle L, (\bar{\cdot}), R_{\mathsf{Ag}}, nom \rangle$ *be an argumentation system for* Ag $\in \{Prop, Op\}$ *such that* $R_{\mathsf{Ag}} = R_s \cup R_{def}^{\mathsf{Ag}}$, *where* R_{def}^{Ag} *is* Ag*'s defeasible rules. An enthymeme dialogue* $d = [m_0, \ldots, m_\ell]$ *between* $Prop$ *and* Op *is said to be* ***well-formed*** *if* $m_i \neq m_j$ *for all* $i \neq j$, *and for all* $i \leq \ell$:

1. $\mathtt{s}(m_i) = Prop$ *if* i *is even, otherwise* $\mathtt{s}(m_i) = Op$;
2. *if* $i = 0$, *then* $\mathtt{l}(m_i) =$ assert, $\mathtt{c}(m_i) \in A^*$, *and* $\mathtt{t}(m_i) = \emptyset$;

3. *if* $i > 0$ *and* $\mathtt{l}(m_i) = \text{assert}$*, then* $\mathtt{c}(m_i) \in E^*$ *and* $\mathtt{l}(\mathtt{t}(m_i)) = \text{assert}$*;*
4. *if* $\mathtt{l}(m_i) = \text{w.d.y.i.}$*, then* $\mathtt{l}(\mathtt{re}(m_i)) = \text{assert}$*;*
5. *if* $\mathtt{l}(m_i) = \text{intend}$*, then* $\mathtt{l}(\mathtt{re}(m_i)) = \text{w.d.y.i.}$ *and* $\mathtt{c}(\mathtt{re}(m_j)) \leq \mathtt{c}(m_i)$*;*
6. *if* $\mathtt{l}(m_i) = \text{w.d.y.a.b.}$*, then* $\mathtt{t}(m_i) = (m_j, \mathtt{t}(m_j))$ *such that* $\mathtt{l}(m_j) = \text{assert}$*, and* $\mathtt{c}(m_i) = \mathtt{c}(\mathtt{t}(m_j))$*;*
7. *if* $\mathtt{l}(m_i) = \text{assumed}$*, then* $\mathtt{l}(\mathtt{re}(m_i)) = \text{w.d.y.a.b.}$ *and* $\mathtt{c}(m_i) \in A^*$ *such that* $\mathtt{c}(\mathtt{re}(m_i)) \leq \mathtt{c}(m_i)$*;*
8. *if* $\mathtt{l}(m_i) = \text{meant}$*, then* $\mathtt{l}(\mathtt{re}(m_i)) = \text{assumed}$*,* $\mathtt{c}(m_i) \in A^*$ *and* $\mathtt{c}(m_i) \neq \mathtt{c}(\mathtt{re}(m_i))$*;*
9. *if* $\mathtt{l}(m_i) = \text{agree}$*, then* $\mathtt{l}(\mathtt{re}(m_i)) = \text{assumed}$*,* $\mathtt{c}(m_i) \in A^*$*, and* $\mathtt{c}(m_i) = \mathtt{c}(\mathtt{re}(m_i))$*.*

The ***topic*** *of* d *(denoted* $\text{Topic}(d)$*) is the label of the conclusion of the argument moved in* m_0*. If* $\exists m_i, m_j$ *(*$0 \leq i < j \leq \ell$*) such that* $j = i + 1$ *and* $\mathtt{l}(m_i) = \mathtt{l}(m_j) = \text{stop}$*, then* $j = l$*. We say that* d *is* ***terminated*** *iff* $\mathtt{l}(m_\ell) = \mathtt{l}(m_{\ell-1}) = \text{stop}$*.*

Step	**Dialogue** d	$\text{IntArg}(m_i)$	$\text{IntTarArg}(m_i)$
1	$m_0 = (Prop, \text{assert}, A, \emptyset, \emptyset)$	A	$\emptyset$
2	$m_1 = (Op, \text{assert}, X', \emptyset, 0)$	X	A
3	$m_2 = (Prop, \text{assert}, B', \emptyset, 1)$	B	C
4	$m_3 = (Op, \text{w.d.y.i.}, \emptyset, 2, \emptyset)$	–	–
5	$m_4 = (Prop, \text{intend}, B, 3, \emptyset)$	B	–
6	$m_5 = (Op, \text{w.d.y.a.b.}, X', \emptyset, (2, 1))$	–	–
7	$m_6 = (Prop, \text{assumed}, C, 5, \emptyset)$	–	–
8	$m_7 = (Op, \text{meant}, X, 6, \emptyset)$	X	–

Table 2
The enthymeme dialogue d from Example 1. The internal structure of arguments and enthymemes moved in d is shown in Figure 2.

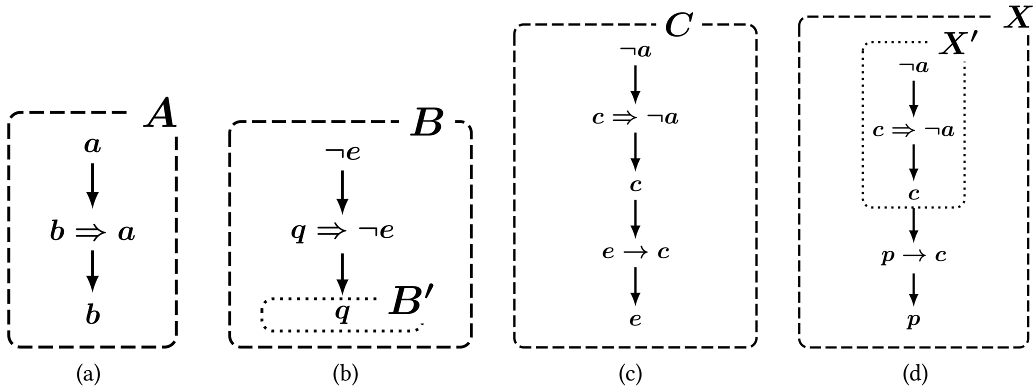

Figure 2: The internal structure of arguments (enclosed by a dashed line) and enthymemes (enclosed by a dotted line) moved in d shown in Table 2.

Henceforth, we use ‘dialogue’ as shorthand for a ‘well-formed enthymeme dialogue’. A dialogue d’s *dialogue framework* is a 4-tuple $DF_d = \langle M^{YN}, T, Re, S \rangle$ where $M^{YN} = M \cup$

$\{\textsc{Yes}, \textsc{No}\}$, M is the set of moves in d (excluding stop moves), Yes and No are auxiliary elements used to respectively confirm and reject the other participant's interpretation of an enthymeme, T is a *defeat* relation, Re a *reply* relation, and S a *support* relation. Recall that a dialogue framework is used to determine the dialogical status of the dialogue moves, similarly to a Dung AF which is used to determine the acceptability status of arguments.

When a move is made in d, it is added to M^{YN} (unless it is a stop). An agent Ag's assert move m_i moves the content of m_i as a defeat on the content of m_i's target, whereas Ag makes a w.d.y.a.b. move to question the defeat relation between two moves m_j and m_k ((m_j, m_k) is the target of m_i). So, in both cases, we add a defeat relation between m_i and its target; in all other cases T remains the same.

If m_i is a w.d.y.i., or intend, or assumed that replies to m_j, then this reply relation is added to Re, since these moves respectively request the intended argument of a move, provide the intended argument of a move and provide an interpretation of an enthymeme. If m_i is a meant move, a reply relation between No and m_j is added to Re. Essentially, m_i has two effects on DF_d: firstly, the sender Ag of m_i rejects its interlocutor's Ag′'s understanding of Ag's enthymeme E, thus, we interpret m_i as a negative reply to m_j; secondly, using m_i, Ag reveals the intended argument from which E was constructed. If we are to draw accurate conclusions regarding the acceptability of moves in DF_d, then m_i is not enough to express both effects. Hence, we use No to validate that Ag′'s interpretation of E is not acceptable, and (as we will see below) we connect m_i (through a support relation) to the move m_k, which contains E, in order to show that the acceptability of m_k (i.e., E) influences the acceptability of m_i (i.e., the intended argument from which E was constructed). In this way, changes in the status of m_i do not influence the status of m_j which remains unacceptable in DF_d since Ag′'s understanding of E was mistaken. In all other cases Re does not change.

If m_i is an intend move with content A, then a support relation between m_i and m_j is added to S. Intuitively, prior to m_i, m_k replies to (and so challenges) m_j, by requesting the sender of m_j to provide the intended argument from which the enthymeme E of m_j was constructed (i.e., A extends E). m_i satisfies this request (and so replies to m_k), and thus m_i is moved to support m_j against m_k's challenge (essentially m_k questions why m_j was moved, whereas m_i justifies why m_j was moved). A meant or agree move m_i also provides the intended argument of a move m_j (where m_j is targeted by a move m_k, the defeat relation (m_k, m_j) is challenged by a move m_h requesting the interpretation of m_j's content by the sender Ag of m_k, a move m_g replies to m_h providing Ag's interpretation of m_j's content and m_i replies to m_g) and so we add a support relation between m_i and m_j. Finally, similarly to a meant move, an agree move has two effects on DF_d: by moving m_i the sender Ag of m_i confirms Ag′'s (i.e., the other participant's) understanding of Ag's enthymeme E and, thus, we interpret m_i as a positive reply to m_g, thus supporting it; secondly, using m_i, Ag reveals the intended argument from which E was constructed. Hence, we use Yes to validate that Ag′'s interpretation of E is acceptable, and we connect m_i (through a support relation) to the move m_j, which contains E, in order to show that the acceptability of m_j (i.e., E) influences the acceptability of m_i (i.e., the intended argument from which E was constructed). In this way, changes in the status of m_i do not influence the status of m_g which remains acceptable in DF_d since Ag′'s understanding of E was correct. In all other cases S remains the same. For an example of a DF see Figure 3.

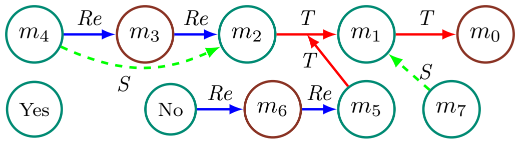

Figure 3: The DF_d instantiated by d in Table 2. Each arrow represents a relation between moves in the DF_d. There is a single complete (and hence preferred, grounded and stable) labelling L of the DF_d. IN and OUT moves are enclosed by a green, respectively brown, circle.

Definition 5. *Let* $DF_{d_0} = \langle M_0^{YN}, T_0, Re_0, S_0 \rangle$ *be the dialogue framework of a dialogue* $d_0 = [m_0]$ *where* $DF_{d_0} = \langle m_0, \emptyset, \emptyset, \emptyset \rangle$. *Let* $d_\ell = [m_0, \dots, m_\ell]$ *expand* d_0. *Then, for each* $DF_{d_i} = \langle M_i^{YN}, T_i, Re_i, S_i \rangle$, *where* $0 \leq i \leq \ell$ *we have that:*

1. $M_i^{YN} = M_{i-1}^{YN} \cup \{m_i\}$ *if* $\texttt{l}(m_i) \neq \mathsf{stop}$, *otherwise* $M_i^{YN} = M_{i-1}^{YN}$;
2. $T_i = T_{i-1} \cup \{(m_i, \texttt{t}(m_i))\}$ *if* $\texttt{l}(m_i) \in \{\mathsf{assert}, \mathsf{w.d.y.a.b.}\}$ *and* $\texttt{t}(m_i) \neq \emptyset$, *otherwise* $T_i = T_{i-1}$;
3. $Re_i = Re_{i-1} \cup \{(m_i, \texttt{re}(m_i))\}$ *if* $\texttt{l}(m_i) \in \{\mathsf{w.d.y.i.}, \mathsf{intend}, \mathsf{assumed}\}$, $Re_i = Re_{i-1} \cup \{(No, \texttt{re}(m_i))\}$ *if* $\texttt{l}(m_i) = \mathsf{meant}$, *otherwise* $Re_i = Re_{i-1}$;
4. $S_i = S_{i-1} \cup \{(m_i, \texttt{re}(\texttt{re}(m_i)))\}$ *if* $\texttt{l}(m_i) = \mathsf{intend}$, *else* $S_i = S_{i-1} \cup \{(m_i, m_j)\}$ *if* $\texttt{l}(m_i) = \mathsf{meant}$ *and* $\texttt{t}(\texttt{re}(\texttt{re}(m_i))) = (m_k, m_j)$, *else* $S_i = S_{i-1} \cup \{(Yes, \texttt{re}(m_i)), (m_i, m_j)\}$ *if* $\texttt{l}(m_i) = \mathsf{agree}$ *and* $\texttt{t}(\texttt{re}(\texttt{re}(m_i))) = (m_k, m_j)$, *otherwise* $S_i = S_{i-1}$, *where* $0 \leq j < k < i$.

We define a *complete labelling* on DF_d of a dialogue d as a function assigning IN to a move m_i in d iff: 1) every move replying to m_i is OUT; 2) every move m_j targeting m_i, and such that the defeat relation from m_j's enthymeme to m_i's enthymeme is not targeted by an IN move, is OUT, and; 3) every move that is supported by m_i is IN. A move m_i is assigned OUT iff either a move replying to m_i is IN, or a move targeting m_i (such that the defeat relation is not targeted by an IN move) is IN, or a move that is supported by m_i is OUT (intuitively, the content E of a move m_j supported by m_i is an enthymeme of m_i's content and so m_j's acceptability influences m_i's acceptability); a move is UNDEC iff it is neither IN nor OUT.

Definition 6. *Let* $DF_d = \langle M^{YN}, T, Re, S \rangle$. *We define a* ***complete labelling*** *on* DF_d *to be a function* $\mathsf{L} : M^{YN} \to \{$IN, OUT, UNDEC$\}$ *such that, for every* $m_i \in M^{YN}$:

1. $\mathsf{L}(m_i) =$ IN *iff for all* $m_j, m_k \in M^{YN}$:
   (a) *if* $(m_j, m_i) \in Re$ *then* $\mathsf{L}(m_j) =$ OUT, *and*
   (b) *if* $(m_j, m_i) \in T$, *and* $\nexists m_k \in M^{YN}$ *such that* $\texttt{t}(m_k) = (m_j, m_i)$ *and* $\mathsf{L}(m_k) =$ IN, *then* $\mathsf{L}(m_j) =$ OUT, *and*
   (c) *if* $(m_i, m_j) \in S$, *then* $\mathsf{L}(m_j) =$ IN;
2. $\mathsf{L}(m_i) =$ OUT *iff there is some* $m_j, m_k \in M^{YN}$ *such that:*
   (a) $(m_j, m_i) \in Re$ *and* $\mathsf{L}(m_j) =$ IN, *or*
   (b) $(m_j, m_i) \in T$, *and* $\nexists m_k \in M^{YN}$ *such that* $\texttt{t}(m_k) = (m_j, m_i)$ *and* $\mathsf{L}(m_k) =$ IN, *and* $\mathsf{L}(m_j) =$ IN, *or*
   (c) $(m_i, m_j) \in S$ *and* $\mathsf{L}(m_j) =$ OUT.

If L is a complete labelling on a DF_d, then $\mathsf{in}_d(\mathsf{L})$, $\mathsf{out}_d(\mathsf{L})$ and $\mathsf{undec}_d(\mathsf{L})$ respectively denote the set of all moves labelled IN, OUT and UNDEC. Below we also define *preferred, preferred* and *stable labellings* on DF_d (for an example see Figure 3).

Definition 7. *Let DF_d be the dialogue framework of a dialogue d and let L be a complete labelling function on DF_d. Then:*

1. L *is a **preferred labelling** function on DF_d iff there does not exist a complete labelling function L' on DF_d such that $\mathsf{in}_d(\mathsf{L}) \subset \mathsf{in}_d(\mathsf{L}')$;*
2. L *is the (unique) **grounded labelling** function on DF_d iff there does not exist a complete labelling function L' on DF_d such that $\mathsf{in}_d(\mathsf{L}') \subset \mathsf{in}_d(\mathsf{L})$;*
3. L *is a **stable labelling** function on DF_d iff $\mathsf{undec}_d(\mathsf{L}) = \emptyset$.*

We define an *argumentation theory instantiated by a dialogue* d as a tuple $AT_d = \langle AS_d, K_d \rangle$, where the logical language L, the contrariness function $(\bar{\cdot})$, the naming function nom, and the strict rules R_s of AS_d are those shared by the participants of d. The *defeasible rules* of AS_d is the set of the defeasible rules revealed during d, excluding the inference rules of enthymemes in assumed moves as these are assumptions which may be mistaken (in case they are correct, they are repeated within the content of an agree move following an assumed move). A premise ϕ of an enthymeme E belongs to the set of *premises* in AT_d (i.e., K_d) iff ϕ is a leaf in E and if the intended argument A from which E was constructed has been revealed by the sender of E, then ϕ is also a leaf in A.

Definition 8. *Let $d = [m_0, \ldots, m_\ell]$ be a dialogue between Prop and Op, where $AS_{\mathsf{Ag}} = \langle L, (\bar{\cdot}), R_{\mathsf{Ag}}, nom \rangle$ is the argumentation system for $\mathsf{Ag} \in \{Prop, Op\}$, K_{Ag} is the knowledge base of Ag and $R_{\mathsf{Ag}} = R_s \cup R_{def}^{\mathsf{Ag}}$. Let $DF_d = \langle M^{YN}, T, Re, S \rangle$ be the dialogue framework of d. The* **set of defeasible rules in** *d is $\mathsf{DefDRules}(d) =$*

$$\bigcup_{i=0}^{\ell} \left\{ \mathsf{lab}_{\mathsf{c}(m_i)}(n) \in R^*_{def} \mid \mathtt{l}(m_i) \neq \mathsf{assumed}, \mathsf{c}(m_i) \in E^* \textit{ and } n \in \mathsf{Nodes}(\mathsf{c}(m_i)) \right\}.$$

The **set of premises of** Ag **in** *d is denoted $\mathsf{DPrem}_{\mathsf{Ag}}(d) =$*

$$\bigcup_{i=0}^{l} \left\{ \mathsf{lab}_{\mathsf{c}(m_i)}(n) \;\middle|\; \begin{array}{c} \mathsf{s}(m_i) = \mathsf{Ag}, \mathtt{l}(m_i) \notin \{\mathsf{w.d.y.i.}, \mathsf{w.d.y.a.b.}, \mathsf{assumed}\}, \\ n \in \\ \mathsf{Leaves}(\mathsf{c}(m_i)), \textit{ and } \nexists m_j \ (j \leq \ell) \textit{ such that } (m_j, m_i) \in S \\ \textit{and } \ n \notin \mathsf{Leaves}(\mathsf{c}(m_j)) \end{array} \right\}.$$

The **set of all premises in** *d is denoted $\mathsf{DPrem}(d) = \mathsf{DPrem}_{Prop}(d) \cup \mathsf{DPrem}_{Op}(d)$.*

The **argumentation theory instantiated by a dialogue** *d is $AT_d = \langle AS_d, K_d \rangle$, where $AS_d = \langle L, (\bar{\cdot}), R_s \cup \mathsf{DefDRules}(d), nom \rangle$ and $K_d = \mathsf{DPrem}(d)$.*

The **argumentation theory of** Ag **instantiated by a dialogue** *d is $AT_d^{\mathsf{Ag}} = \langle AS_d^{\mathsf{Ag}}, K_d^{\mathsf{Ag}} \rangle$, where $AS_d^{\mathsf{Ag}} = \langle L, (\bar{\cdot}), R_{\mathsf{Ag}} \cup \mathsf{DefDRules}(d), nom \rangle$ and $K_d^{\mathsf{Ag}} = \mathsf{DPrem}(d) \cup K_{\mathsf{Ag}}$.*

4. Discussion

If a dialogue d is *exhaustive* and its participants *honest*, the dialectical status of the moves in the DF_d (determined by a complete, preferred, grounded and stable labelling) is sound and complete with respect to the dialectical status of the arguments in the AF instantiated by the contents of the moves made in the dialogue (determined by a complete, preferred, grounded, and stable labelling, resp.). Essentially, if d's participants are *honest*, they only assert an enthymeme constructed from the intended argument of their move and such that it does defeat the intended targeted argument (i.e., what the sender assumes to be the target move's intended argument) according to the $ASPIC^+$ definition of defeat. It also means that whenever they reveal their understanding of the argument from which their counterpart's enthymeme E is constructed, the argument does indeed extends E, and whenever they reveal the intended argument of their own move, this is indeed the argument they intended. If d is *exhaustive*, then any available move that can be made, is indeed made. In other words, if a participant can move an argument (constructed by the contents of the moves made in the dialogue) as a defeat against the content of another move made in the dialogue, or question a defeat relation or reply to a move, they indeed do so. Due to lack of space we cannot present here the technical details for our claim, but for a better understanding of our claim the reader can look at Example 2.

Example 2. *Figures 4.(a) and 4.(b) respectively show the DF_d of Table 2's exhaustive and honest dialogue d [2], and the AF_{AT_d} instantiated by AT_d. We also show a valid complete (resp. preferred, grounded and stable) labelling of DF_d's moves given by the labelling function L on DF_d, and a valid complete (resp. preferred, grounded and stable) labelling of arguments in AF_{AT_d} given by the labelling function L' on AF_{AT_d}. In both frameworks the topic a of d is labelled OUT, (since m_0 is labelled OUT in DF_d and $\mathsf{IntArg}(m_0)$ is labelled OUT in AF_{AT_d}) which means that a does not belong to the complete (resp. preferred, grounded and stable) extension [13]. Moreover, the acceptability of moves in d coincides with the acceptability of their intended arguments in the AF_{AT_d} instantiated by AT_d.*

This paper introduces a novel dialogue system that allows participants to seek and provide clarifications regarding possible misinterpretations that arise from the use of enthymemes, and instantiates a dialogue framework that is used to determine the dialogical status of the dialogue moves. As far as we are aware, there are no other works that handle misunderstandings that may occur between the participants due to the use of enthymemes and also propose a mechanism for determining the status of the dialogue moves. Our work paves the way for a sound and complete dialogue system that accommodates use of any kind of enthymeme, and where any kind of uncertainty that arises from their use can be resolved. This is important since it ensures that the dialogue can be played out such that an enthymeme moved in the dialogue is only justified in the case that its intended argument is justified by the contents of the moves made in the dialogue. In future work we will explore how enthymemes may be used to give a strategic advantage to a participant.

[2] We assume that C and X are preferred to A since owing money takes precedence over eating at a restaurant, and B is preferred to C' since C' is just a statement.

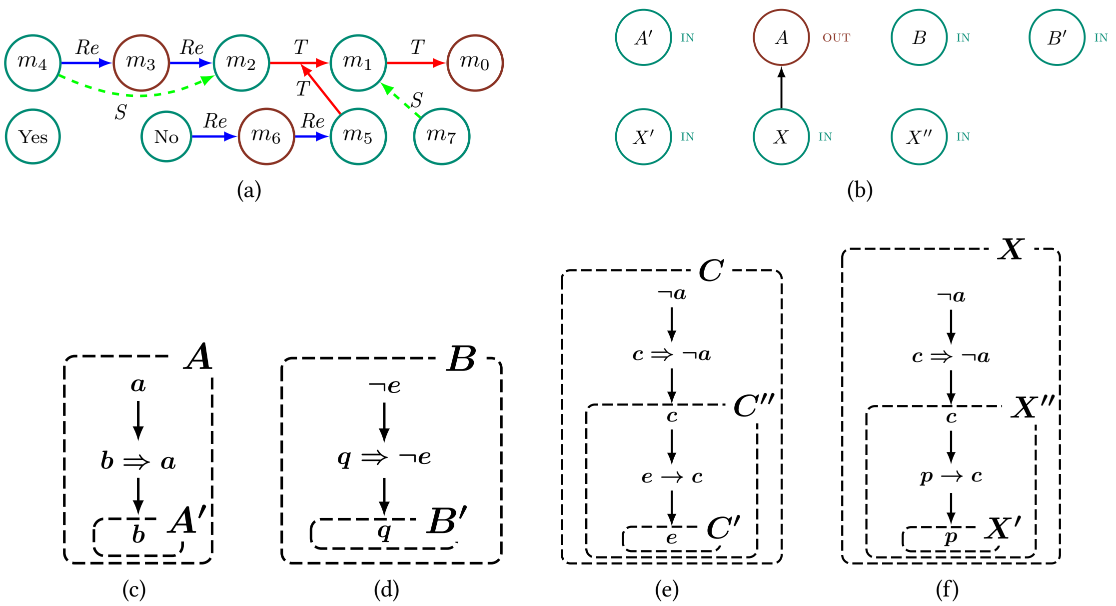

Figure 4: (a) The unique complete (preferred, grounded and stable) labelling L on the DF_d instantiated by d in Table 2 (arrows represent relations between moves in DF_d). IN and OUT moves are respectively enclosed by green and brown circles. (b) The AF_{AT_d} instantiated by AT_d. A complete (resp. preferred, grounded, stable) labelling function L′ on AF_{AT_d} is shown. IN and OUT arguments are enclosed by a green and a brown circle, respectively. Note, this is the only possible complete (resp. preferred, grounded, stable) labelling on AF_{AT_d} for this example. (c)-(f) The arguments instantiated by AT_d enclosed by a dashed line.

References

[1] P. Besnard, A. Garcia, A. Hunter, S. Modgil, H. Prakken, G. Simari, F. Toni, Introduction to structured argumentation, Journal of Argument & Computation (2014) 1–4.

[2] W. Douglas, Informal logic: A handbook for critical argument, Cambridge University Press, 1989.

[3] S. Modgil, Revisiting abstract argumentation frameworks, in: Proceedings of Theory and Applications of Formal Argumentation, 2013, pp. 1–15.

[4] E. Black, A. Hunter, A relevance-theoretic framework for constructing and deconstructing enthymemes, Logic and Computation (2012) 55–78.

[5] A. Hunter, Real arguments are approximate arguments, in: Proceedings of AAAI Conference on Artificial Intelligence, 2007, pp. 66–71.

[6] D. Walton, C. Reed, Argumentation schemes and enthymemes, Synthese (2005) 339–370.

[7] F. D. de Saint-Cyr, Handling enthymemes in time-limited persuasion dialogs, in: Proceedings of International Conference on Scalable Uncertainty Management, 2011, pp. 149–162.

[8] A. Xydis, C. Hampson, S. Modgil, E. Black, Enthymemes in dialogues, in: Proceedings of Computational Models of Argument, 2020, pp. 395–402.

[9] H. Prakken, Coherence and flexibility in dialogue games for argumentation, Logic and Computation (2005) 1009–1040.
[10] A. Xydis, C. Hampson, S. Modgil, E. Black, Towards a sound and complete dialogue system for handling enthymemes, in: International Conference on Logic and Argumentation, 2021, pp. 437–456.
[11] S. Modgil, H. Prakken, Abstract rule-based argumentation, Handbook of Formal Argumentation (2018) 286–361.
[12] P. M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, Artificial Intelligence (1995) 321–357.
[13] M. Caminada, On the issue of reinstatement in argumentation, in: European Workshop on Logics in Artificial Intelligence, 2006, pp. 111–123.
[14] E. Black, A. Hunter, Using enthymemes in an inquiry dialogue system, in: Proceedings of International Conference on Autonomous Agents and Multi-agent Systems, 2008, pp. 437–444.
[15] S. A. Hosseini, Dialogues incorporating enthymemes and modelling of other agents' beliefs, Ph.D. thesis, King's College London, 2017.