A formal analysis of enthymematic arguments

Sjur K. Dyrkolbotn[1], **Truls Pedersen**[2],
[1] Western Norway University of Applied Sciences
[2] University of Bergen
sdy@hvl.no, truls.pedersen@infomedia.uib.no

Abstract

We provide a simple formalisation of enthymematic arguments, based on formal argumentation theory. We start from a simple representation of arguments as sequences of formulas and rules. Regular arguments are those that explicitly lists the conclusions of all rules applied, while also establishing every premise of every rule used. An enthymematic argument is then defined as a sequence that does not satisfy this property, but which can be extended to such a sequence in one or more ways. Borrowing terminology from the informal logic literature on enthymematic arguments, we then define the "crater" of an enthymematic argument as a set of arguments, namely those that minimally extend the enthymematic argument in the appropriate way. We go on to propose a notion of attack between enthymematic arguments, allowing us to represent them as nodes in a Dung-style attack graph. We also prove a characterisation result, providing necessary and sufficient conditions for the acceptance of an enthymematic argument under Dung-style semantics based on admissible sets.

1 Introduction

If I argue that I am hungry, so I should go to the store, you are unlikely to write me off as irrational. What I said makes sense, on the assumption that I am at home without food. It also makes sense if I am at home and there *is* some food, as long as the food I have is not food I want to eat. Similarly, what I said makes sense if I am not at home, but at work, if I forgot my lunch. There are countless variations on this theme, of course, providing a reasonable interpretation of what I said. I did not present a complete argument for the conclusion that I should go to the store, but any charitable listener will be able to fill in the gaps and form a meaningful hypothesis about my meaning. This is the typical situation in natural language, when people argue and give reasons for actions. Complete specifications are not feasible, but incomplete approximations are commonplace and easily understood in most cases.

However, trouble can arise in case of misaligned interpretations. If my Google AI adviser hears my argument for going to the shop and says "no, it's not your lunch break yet, so you should not leave your office", my assessment of this counterargument depends rather crucially on whether or not I am at home with my kids or at work with my colleagues. Still, if I have not informed my adviser either way, I can hardly blame her for assuming the worst and warning me accordingly. Indeed, in a case like this, it seems clear that if I am misinterpreted, the burden to defend my conclusion falls on me. The AI adviser makes a case against leaving the office. What it said will attack any argument for going to the shop that depends on the fact that I may leave my office.

In other words, the AI adviser attacks some interpretations of my argument. Since these are interpretations I did not rule out, the attack succeeds also as an attack on my underspecified argument. This, in essence, is how we conceive of enthymematic arguments in this article, as partially explicated arguments A, B such that A attacks B if *some interpretation* of A attacks *some interpretation* of B. We formalise this in the following, culminating in a Dung-style argumentation semantics for enthymematic arguments that also leads to a simple and natural characterisation of which such arguments we can accept.

The structure of the paper is as follows. In Section 2 we describe some of the philosophical history of the notion of enthymemes and the importance of them in relation to artificial intelligence. In Section 3 we consider the question of how enthymematic arguments should be defined, proposing a definition based on the theory of structured argumentation. In Section 4 we define and discuss semantics for enthymematic arguments in terms of abstract argumentation frameworks. In Section 5 we present the main result of the paper, providing necessary and sufficient conditions for the acceptance of enthymematic arguments. In Section 6 we offer a short conclusion and directions for future work.

2 Enthymemes

The concept of an *enthymeme* was first discussed in Aristotle's studies of logic and rhetoric. It is meant to capture what is "left in the mind" after an argument has been put forward (from Greek: *en-* "in" and *thymos* "mind"). Today, the term is often associated with rhetoric, but it has also been studied in argumentation theory (see e.g., [Walton, 2008]). Enthymemes are informally described in various ways, such as [Gilbert, 1991]:

> "Everyone agrees that an enthymeme is an argument. Most writers also agree that enthymemes, even though they are formally invalid, are not bad arguments simply as a result of being enthymematic, but rather lack something that non-enthymemes do not."

What enthymemes lack that non-enthymemes do not vary according to different authors. Premises are particularly often said to be lacking, either because they are "unexpressed", "suppressed", "implicit", "hidden", or "unstated". These nuances aside, an enthymeme may also lack other elements than premises, such as the conclusion or even the rules applied. Such a lack may be similar to lacking part of what would constitute "warrant" in Toulmin's framework [Toulmin, 2003]. Not everybody agrees that enthymemes are arguments. In [Goddu, 2016] it is argued that lacking something essential is essential to enthymemes, making them incompatible with what it means to be an argument.

Metaphysical questions aside, enthymemic arguments are interesting for many reasons. To us, their potential application in artificial intelligence is of particular interest. Here, formalising enthymemes can provide a natural approach to reasoning about agents who are unable or unwilling to provide complete arguments and explanations for why they believe certain things or act in certain ways. A general inability to completely specify one's point of view seems endemic to complex reasoners, so an ability to deal with incomplete specifications seems like a crucial feature for complex social agents. Furthermore, efficiency gains can be made by taking an economic approach to reasoning and explanation, relying on our ability to process enthymematic arguments rather than asking always for the most "complete" picture possible.

We should clarify at the outset that there is some disagreement about whether an enthymeme is an argument that is (i) implicitly describing premises, (ii) missing premises (iii) missing premises or a conclusion, or (iv) missing something less specific (e.g., lacking in clarity or precision). We take the position here that enthymemes may be missing premises or explicit references to inference rules. However, we insist that the conclusion is at least implicitly provided. There are technical and philosophical reasons for this assumption. The technical reasons will appear when we model enthymemes in ways compatible with ASPIC$^+$and related argumentation frameworks [Modgil and Prakken, 2014]. In these frameworks, explicitly refering to the inference rules which the arguer applies in her argument exposes the argument for *undercutting* arguments.

One philosophical reason is that for the kind of arguments we are interested in trying to capture, it is a reasonable expectation that the arguer makes the conclusion known to his audience. We believe the arguer has an obligation to make the conclusion known in order to make the proposed argument honestly open to refutation. If we permit missing conclusions it may become unreasonably difficult to attack a proposal, since a chain of "that is not what I meant"-defences can then be extended indefinitely. We believe these considerations are important particularly in advanced AI-agents [Rahwan and Simari, 2009], as we will allude to in the running example. While we do not deny that there are cases in which conclusions *ought* or *must* be left underspecified, these cases are not the subject of our investigation. We believe they have more to do with *ambiguity* than with *incompleteness*, and we prefer to keep these distinct sources of uncertainty apart in the formal analysis.

3 Defining enthymematic arguments

We assume that we are working with a propositional language $\mathscr{L}$, reasoning about formulas from this language using a set of strict rules S and a set of defeasible rules $\mathscr{D}$. We assume that every rule $r_i \in \mathsf{S} \cup \mathscr{D}$ has the form $r_i = (P_i, c_i)$ where $P_i \subseteq \mathscr{L}$ is the set of premises of r_i and $c_i \in \mathscr{L}$ is its conclusion. We introduce some simple projections for the premises and conclusions of rules, and permit them also to apply to formulas, mapping $\mathscr{D} \cup \mathsf{S} \cup \mathscr{L} \to 2^{\mathscr{L}}$ and $\mathscr{D} \cup \mathsf{S} \cup \mathscr{L} \to \mathscr{L}$, respectively:

$$P(x) = \begin{cases} P_x & \text{if } x = (P_x, c_x) \\ \{x\} & otherwise \end{cases} \quad c(x) = \begin{cases} c_x & \text{if } x = (P_x, c_x) \\ x & otherwise \end{cases}$$

Example 1. Consider the example from the introduction. Here are some propositions denoting the claims involved, as well as some additional claims that we will use to fill in the "gaps" of our argument:

p_1:	"I should go to the store"	p_5:	"I am hungry"
p_2:	"Fresh food in the fridge"	p_6:	"I am at home"
p_3:	"I have food I want to eat"	p_7:	"I am at work"
p_4:	"I forgot my lunch"	p_8:	"Lunch break"

The rules we rely on are the following:

$$\frac{p_3,\ \neg p_5}{p_1}\, r_1, \quad \frac{p_7,\ \neg p_8}{\neg n(r_5)}\, r_2, \quad \frac{p_3,\ p_5}{\neg p_1}\, r_3$$
$$\frac{p_2,\ \neg p_6}{\neg p_3}\, r_4, \quad \frac{p_4,\ p_7}{\neg p_3}\, r_5$$

The enthymeme the reader hypothetically accepted in the introduction is $E^\star = (p_5, p_1)$: "I am hungry, so I (should) go to the store". We are not mentioning the rule(s) we are applying. Indeed, from the little information we have expressed we are not even able to apply any rules. These utterances are sufficient for the recipient to "fill in the gaps". Suppose our theory includes the observations that $\neg p_2$: "(There is no) fresh food in the fridge", p_4: "I forgot my lunch", p_5: "I am hungry", and p_7: "I am at work". Then the theory supports the application of r_5 yielding $\neg p_3$. Together with p_5 from the theory, we may now apply r_1 and obtain p_1. The enthymeme we stated can be expanded by making these observations explicit to form, for example, the explicit argument $E_1 = (p_5, p_4, p_7, r_5, \neg p_3, p_5, r_1, p_1)$.

As is usual in the theory of structured argumentation, we rely on the definition of contrary formulas, $\overline{} : \mathscr{L} \to 2^{\mathscr{L}}$, denoting the set of formulas that is *contrary* to a given formula.

Since the order of the reasoning rules we invoke can be significant when evaluating enthymematic arguments, we will represent arguments as sequences of rules rather than as proof trees. Furthermore, we will also be interested in arguments that include redundant premises and rules, so we will not stipulate that the rules occurring in an argument are all needed to establish premises required by subsequent rules.

In order to accomodate undercutting attacks, we name all defeasible rules by a naming function similar to how ASPIC$^+$ models this. For uniformity, we define the function $n : \mathscr{D} \cup \mathsf{S} \cup \mathscr{L} \to \mathscr{L}$, where we require that

- $n(x) \in \mathscr{L}$ if $x \in \mathscr{D}$,
- $n(x) = \top$ if $x \in \mathsf{S}$, and
- $n(x) = x$ if $x \in \mathscr{L}$.

This means that every defeasibly rule is named, strict rules have a vacuous name and every formula names itself. This notion is altered slightly from ASPIC$^+$'s terminology, but not substantially.

Formally speaking, any argument A in our formalism will be instantiated by a sequence $A = (x_1, x_2, \ldots, x_n)$ of rules and formulas, satisfying the constraints in Definition 1.

Definition 1. Given a theory T

- An *argument* based on T (using $\mathscr{D}$ and S) is a sequence $A = (x_1, x_2, \ldots, x_n)$ such that $\forall 1 \leq i \leq n : x_i \in \mathscr{D} \cup \mathsf{S} \cup \mathscr{L}$.
- An argument $A = (x_1, x_2, \ldots, x_n)$ is said to be *complete* if
  - $\forall 1 \leq i \leq n : \begin{cases} x_i \in \mathsf{T}, \text{ or} \\ \forall \phi \in P(x_i) : \phi \in \{c(x_j) \mid 1 \leq j < i\} \end{cases}$
    and
  - $\forall 1 \leq i \leq n : \exists i \leq j \leq n : c(x_i) = x_j$.

The conditions on complete argument can be intuitively justified as follows. The first condition requires that an element is either the explication of a formula in the theory, or that all formulas/premises this element relies on has already been established earlier in the sequence. The second requirement states that the conclusion of every rule must be explicated at some point after the rule has been applied.

We define the conclusion of $A = (x_1, x_2, \ldots, x_n)$ as $c(A) = c(x_n)$. That is, if A ends with a rule then the conclusion of A is the conclusion of the final rule applied in A. If, on the other hand, A ends with a formula, the conclusion of A is this formula. The set of all arguments based on T is denoted by $\overline{\mathsf{A}}$, while the set of complete arguments is denoted by A.

When we do not need to reference individual rules, we generally use upper-case letters like A, B, C etc. to denote arguments. However, any such abstract argument corresponds to an actual sequence of rules meeting the requirements from Definition 1.

Example 2 (Example 1 continued)**.** Continuing from the previous example, it is easy to verify that the uttered $E^\star$ complies with the weakly constrained definition of an argument: every element in the sequence is either a rule or a formula. It is *not* a complete argument, however. We have $P(p_5) = \{p_5\} \subseteq \mathsf{T}$, but $P(p_1) = \{p_1\}$, and p_1 is neither in the theory nor is it (the conclusion of) an earlier element.

After we filled in the gaps to obtain $E_1 = (p_5, p_4, p_7, r_5, \neg p_3, r_1, p_1)$ we obtained an argument satisfying every condition used to characterise a *complete* argument. Every element in the sequence is such that, if it is a formula, then it is the conclusion of a previously applied rule or already a part of the theory. Also, every element x_i satisfies the condition that it or some later element x_j affirms the conclusion of it. If it is a formula, then it affirms itself. This forces the conclusions of the applied rules to be explicitly listed in the argument after they are derived. Finally, all of $P(r_5)$ and $P(r_1)$ occur in the sequence before the respective rules occur.

Suppose we added p_2 to E_1 in Example 4. We get a new argument $E_1' = (p_2, p_4, p_7, r_5, \neg p_3, p_5, r_1, p_1)$ which is identical to E_1 but with the proposition p_2 appended to it. Since p_2 is in the theory, E_1' is still complete. However, p_2 does not occur as a premise of any of the applied rules, nor is if otherwise connected with the conclusion. Furthermore, E_1' has E_1 as a strict complete subsequence. In what follows, we will generalise this observation to formalise the notion of a *minimally complete* argument. This, in turn, will serve as a basis for our formal definition of what it means to be an enthymematic argument.

For any positive integer n, we let $[n]$ denote the set of natural numbers between 1 and n. A sequence $A = (x_1, x_2, \ldots, x_n)$ can then be conventionally written as $(x_i)_{i \in [n]}$. We say that a function $f : [n] \to [m]$ from sets of positive integers to sets of positive integers is strictly increasing if $f(x) < f(y)$ whenever $x < y$, for $x, y \in [n]$. Then the notion of a subsequence is formally defined by the condition that $A \preceq B$ for arguments $A = (x_1, x_2, \ldots, x_n)$ and $B = (y_1, y_2, \ldots, y_m)$ iff there is a strictly increasing function $f : [n] \to [m]$ such that $\forall i \in [n] : x_i = y_{f(i)}$. That is, $A \preceq B$ if A can be obtained from B by deleting rules or formulas. Equivalently, B is obtained from A by filling in the gaps with rules or formulas.

This points to a formalisation of the intuition we had about the meaning of enthymematic arguments. Specifically, we arrive at the following general definition of the *crater* of an enthymematic argument (see for example [Paglieri and Woods, 2011]).

Definition 2. For all arguments $A \in \overline{\mathsf{A}}$, we say that A is *minimally complete* if there is no complete argument $B \prec A$ such that $c(A) = c(B)$.

Definition 3. Given an argument $A = (x_1, x_2, \ldots, x_n)$ based on T, the *crater* of A, denoted $I(A)$ contains all minimally complete arguments $B = (y_1, y_2, \ldots, y_m)$ such that $c(A) = c(B)$ and either $A \preceq B$ or $B \preceq A$.

Definition 4. For any argument A, we say that A is

- *incoherent* if $I(A) = \emptyset$.
- *enthymematic* if $\exists B \in I(A) : A \prec B$,
- *regular* if $I(A) = \{A\}$
- *superfluous* if $\exists B \in I(A) : B \prec A$.

That is, A is an enthymematic argument if its crater contains a minimally complete argument that extends A. This definition rules out other forms of incompleteness in argumentation, e.g., cases where the expression used to express a rule is ambiguous so that two or more interpretations are possible. At the same time, the definition rules out arguments that *cannot* be completed, as well as complete arguments that contain *redundant* rules or formulas (i.e., arguments that are *superfluous*). Such arguments A have craters that are empty or contain subsequences of A.

As a first step towards unpacking the definition further, we record the following simple claim.

Proposition 1. *An argument is regular if, and only if, it is minimally complete.*

Proof. By Definition 3, $I(A)$ contains all minimally complete B such that $A \preceq B$ or $B \preceq A$. By Definition 2, an argument A is minimally complete if, and only, if there is no complete $B \prec A$ such that $c(A) = c(B)$. Hence, if A is minimally complete, there is no minimally complete $B \prec A$. Furthermore, there is no minimally complete B such that $A \prec B$, since A being minimally complete contradicts any such B being so. Since $A \preceq A$, it follows that $I(A) = \{A\}$ if, and only if, A is minimally complete. □

It is also easy to show that any argument belongs to exactly one of the categories listed in Definition 4. Specifically, this follows as a corollary of the following simple observation.

Proposition 2. *For all $A \in \overline{\mathsf{A}}$, we have:*

- *a)* $\exists B \in I(A) : B \prec A \Rightarrow \forall B \in I(A) : B \prec A$ *and*
- *b)* $\exists B \in I(A) : A \prec B \Rightarrow \forall B \in I(A) : A \prec B$

Proof. For a), assume $B \in I(A)$ with $B \prec A$. Assume towards contradiction that there is some $C \in I(A)$ with $A \preceq C$. By Definition 3, this means that C is minimally complete. Since $\prec$ is transitive, we get $B \prec C$. But by Definition 2, this contradicts the fact that C is minimally complete. The argument for b) is similar. □

Corollary 1. *Any argument A is exactly one of the following: incoherent, enthymematic, regular, or superfluous.*

Proof. Consider arbitrary $A \in \overline{\mathsf{A}}$. Obviously, A belongs to at least one of the categories defined in Definition 4. Furthermore, the claim that A belongs to only one of these categories is obviously true if A is incoherent or regular. If A is enthymematic, then it is clearly not incoherent or regular. Moreover, it follows by Proposition 2 a) that A is not superfluous either. Similarly, if A is superfluous, it is obivously not incoherent or regular. Furthermore, it follows by Proposition 2 b) that it is not enthymematic either. □

Example 3 (Example 1 continued)**.** Continuing from the previous example, it is easy to verify that $E^\star$ is indeed enthymematic according to Definition 4. First, notice that its crater is $I(E^\star) = \mathbf{E_1}$ where $\mathbf{E_1}$ is a set of minimally complete arguments that contain all permutations of the internal elements of E_1 that still result in an acceptable elaboration of $E^\star$ (the order of the "missing" elements does not matter, as long as we get a minimally complete argument that extends $E^\star$). To illustrate when we can encounter craters with semantically distinct objects, assume that we replace p_6, p_7 by the default rules $r_6 : (\top, p_6), r_7 : (\top, p_7)$. This is a possible encoding of the state of an AI adviser who has defeasible reasons to think both that I am at home and that I am at work (this encodes uncertainty, in argumentative terms). In this case, the crater of $E^\star$ includes variants of both $F_1 = (p_5, p_4, r_7, p_7, r_5, \neg p_4, r_1, p_1)$ and $F_2 = (p_5, r_6, p_6, \neg p_2, r_4, \neg p_3, r_1, p_1)$. This encodes the AI perspective on $E^\star$. Suppose the theory of the AI contains $\neg p_8$ (no lunch break), because the AI knows that it is not lunch time. Then the AI can form the argument $G = (\neg p_8, r_7, p_7, r_2, \neg n(r_5))$. This argument involves the rule r_2, which can be used to undercut r_5 (intuitively, the argument tells me not to think about food at all when it is not my lunch break). It is easy to verify that G is regular, i.e., its crater consists of G itself. In the next section, we define a semantics according to which G also attacks $E^\star$ in this case, since it attacks F_1.

4 Semantics

We let A denote the set of all regular arguments, namely all A such that $I(A) = \{A\}$. The set $\overline{\mathsf{A}}$, meanwhile, denotes all arguments (so that $\mathsf{A} \subseteq \overline{\mathsf{A}}$).

Definition 5. We define two relations of attack as follows:

- For all $A, B \in \mathsf{A}$ we define R such that $(A,B) \in R$ if, and only if,
$$\exists x \in B : c(A) \in \overline{c(x)} \cup \overline{n(x)}$$
- For all $A, B \in \overline{\mathsf{A}}$ we define $\overline{R}$ such that $(A,B) \in \overline{R}$ if, and only if,
  - $I(B) = \emptyset$ or
  - $\exists A' \in I(A) : \exists B' \in I(B) : (A', B') \in R$

It is easy to see that R and $\overline{R}$ agree on the notion of attack for regular arguments.

Proposition 3. *For all $A, B \in \mathsf{A}$, we have $(A,B) \in \overline{R}$ if, and only if $(A,B) \in R$.*

Proof. Since $A, B \in \mathsf{A}$, we have $I(A) = \{A\}$ and $I(B) = \{B\}$. Hence, the claim follows by Definition 5. □

In other words, $R \subseteq \overline{R}$, so that $\overline{R}$ extends the attack relation to enthymematic arguments. Viewing the set $(\mathsf{A}, \overline{R})$ as an abstract argumentation framework, this means that we obtain semantics also for enthymematic arguments.

Definition 6. Assume given a pair (X,R) where $R \subseteq X \times X$. Then we have the following argumentation semantics for (X,R):

Admissible $\mathsf{Adm}(X,R) = \{S \mid \forall A,B \in S : (A,B) \notin R \;\&\; \forall A \in S : \forall B \in R^-(A) : \exists C \in S : (C,B) \in R\}$.

Complete $\mathsf{Com}(X,R) = \{S \in \mathsf{Adm}(X,R) \mid S = \overline{S}\}$ where for all $S \subseteq X$,
$$\overline{S} = S \cup \{A \in X \mid \forall B \in R^-(A) : \exists C \in S : (C,B) \in R\}.$$

Preferred $\mathsf{Pref}(X,R) = \{S \in \mathsf{Adm}(X,R) \mid \forall S' \in \mathsf{Adm}(X,R) : S \not\subset S'\}$.

Grounded $\mathsf{Ground}(X,R) = \{S \in \mathsf{Com}(X,R) \mid \forall S' \in \mathsf{Com}(X,R) : S' \not\subset S\}$.

Proposition 4. *Let $A, B \in \overline{\mathsf{A}}$ and assume there is some $A' \in I(A)$ that attacks B. Then every $A' \in I(A)$ attacks B.*

Proof. First notice that we have $c(A_1) = c(A_2)$ for all $A_1, A_2 \in I(A)$. That is, all $A' \in I(A)$ have the same conclusion. Notice, moreover, that for all $A_1, A_2, B \in \mathsf{A}$, if $c(A_1) = c(A_2)$, then $(A_1, B) \in R \Leftrightarrow (A_2, B) \in R$. This is because the conclusion of A uniquely determines which arguments A attack. From this, the claim follows: if there is some argument in the crater of A that attacks B, then every argument in the crater of A attacks B. □

Example 4 (Example 1 continued)**.** Once again, consider $E^\star$, from the perspective of the AI (such that default rules r_6, r_7 can be used to argue for p_6, p_7 respectively). The crater consists of variants of F_1 and F_2. Since F_1 involves the rule r_5, $G = (\neg p_8, r_7, p_7, r_2, \neg n(r_5))$ attacks F_1, so by Definition 5 it also attack $E^\star$. Notice how adding $\neg p_7$ to the knowledge base of the AI would prevent this attack on $E^\star$. After the additional knowledge is added, G is no longer a minimally complete argument. However, if I extend my enthymematic argument $E^\star$ to another enthymematic argument $F^\star = (p_5, p_6, p_1)$ I achieve the same effect, since now only variants of F_1 is in the crater. This shows how a formal model of enthymematic argumentation will allow us to deal with more economically expressed arguments in a systematic way, accounting for the semantic effects of leaving arguments underspecified.

5 A characterisation result

Definition 7. Given a set $S \subseteq \overline{\mathsf{A}}$, we define the following sets of corresponding arguments.

1. $Ext(S) = \{A \mid \emptyset \subset I(A) \subseteq S\}$.
2. $Res(S) = \bigcup_{A \in S} I(A)$

Hence, $Ext(S)$ collects all arguments whose craters are subsets of S. In general, we may have $S \not\subseteq Ext(S)$, namely if, and only if, ($\star$) there is some $A \in S$ such that $I(A) \not\subseteq S$. However, if $S \subseteq \mathsf{A}$, then $S \subseteq Ext(S)$, since $I(A) = \{A\}$ for all $A \in \mathsf{A}$. In this case, $Ext(S)$ is an extension of S. $Res(S)$, meanwhile, takes $S \in \overline{\mathsf{A}}$ and returns the union of all craters of elements in S. Hence, $Res(S) \subseteq \mathsf{A}$, providing in all cases a projection of S onto the set of regular arguments. Notice, moreover, that we have $S \cap \mathsf{A} \subseteq Res(S)$, since elements of A are their own craters. It should also be noted that ($\star$) is the case if, and only if, $Res(S) \not\subseteq S \cap \mathsf{A}$ (which is equivalent to $Res(S) \neq S \cap \mathsf{A}$).

We can now prove the following characterisation theorem, showing us how to get from admissible sets of arguments in (A, R) to admissible sets of arguments in $(\overline{\mathsf{A}}, \overline{R})$ and vice versa.

Theorem 1. *For all theories* T *and all* (A, R) *and* $(\overline{\mathsf{A}}, \overline{R})$ *based on* T*, we have the following:*

$$\begin{array}{ll} 1) & S \in \mathsf{Adm}(\mathsf{A}, R) \Rightarrow Ext(S) \in \mathsf{Adm}(\overline{\mathsf{A}}, \overline{R}) \\ 2) & S \in \mathsf{Adm}(\overline{\mathsf{A}}, \overline{R}) \Rightarrow Res(S) \in \mathsf{Adm}(\mathsf{A}, R) \end{array}$$

Proof. **1**) Assume that S is an admissible set in (A, R) and consider $Ext(S)$. We have to show that $Ext(S)$ is independent and defends itself in $(\overline{\mathsf{A}}, \overline{R})$.
Independence: Let $A, B \in Ext(S)$ and assume towards contradiction that $(A, B) \in \overline{R}$. By Definition 5 this means that some $A' \in I(A)$ attacks some $B' \in I(B)$. We choose some such A', B'. Since S is admissible, it follows that there is some $C \in S$ such that $(C, A') \in R$. However, since $A \in Ext(S)$ it follows by Definition 7 that $I(A) \subseteq S$. Hence, $A', C \in S$, contradicting independence of S.
Self-defence: Let $(A, B) \in \overline{R}$ for some arbitrary $B \in Ext(S)$. We have to show that $Ext(S)$ attacks A. By Definition 5 and the fact that A attacks B, we know there is some $A' \in I(A)$ that attacks some $B' \in I(B)$. By Definition 7, we know that $I(B) \subseteq S$. Since S is admissible it then follows that there is $C \in S$ such that $(C, A') \in R$. By Definition 7, we have $C \in Ext(S)$. Moreover, by Definition 5, we get $(C, A) \in \overline{R}$. Hence, $Ext(S)$ attacks A as desired.
2) Assume that S is an admissible set in $(\overline{\mathsf{A}}, \overline{R})$. We have to show that $Res(S) = S \cap \mathsf{A}$ is independent and defends itself in (A, R).
Independence: Assume towards contradiction that there is $A, B \in Res(S)$ such that $(A, B) \in R$. By Definition 7 we have $A', B' \in S$ such that $A \in I(A'), B \in I(B')$. It follows by Definition 5 that $(A', B') \in \overline{R}$, contradicting independence of S.
Self-defence: Consider arbitrary $(A, B) \in R$ such that $B \in Res(S)$. By Definition 7 there is some $B' \in S$ such that $B \in I(B')$. Furthermore, by Definition 5 we have $(A, B') \in \overline{R}$. Since S defends itself, there is some $C \in \overline{\mathsf{A}}$ such that $(C, A) \in \overline{R}$. By Definition 5 this means that there is some $C' \in I(C)$ and some $A' \in I(A)$ such that $(C', A') \in \overline{R}$. Since $A \in \mathsf{A}$, we have $I(A) = \{A\}$, which implies $A = A'$. Furthermore, by Definition 7 we have $C' \in Res(S)$. Hence, $Res(S)$ defends B against A. Since (A, B) as arbitrarily chosen, the claim follows. □

Theorem 2. *For all theories* T *and all* (A, R) *and* $(\overline{\mathsf{A}}, \overline{R})$ *based on* T*, if* $\varepsilon \in \{\mathsf{Com}, \mathsf{Pref}, \mathsf{Ground}\}$ *have the following:*

$$\begin{array}{ll} 1) & S \in \varepsilon(\mathsf{A}, R) \Rightarrow Ext(S) \in \varepsilon(\overline{\mathsf{A}}, \overline{R}) \\ 2) & S \in \varepsilon(\overline{\mathsf{A}}, \overline{R}) \Rightarrow Res(S) \in \varepsilon(\mathsf{A}, R) \end{array}$$

Proof. $\varepsilon = \mathsf{Com}$: 1) Assume $S \in \mathsf{Com}(\mathsf{A}, R)$. We have to show $Ext(S) \in \mathsf{Com}(\overline{\mathsf{A}}, \overline{R})$. By Theorem 1 we know that $Ext(S)$ is admissible, so we only have to show that it is complete. Assume towards contradiction that there is some $A \in \overline{\mathsf{A}} \setminus Ext(S)$ such that

$$\forall (B, A) \in \overline{R} : \exists C \in Ext(S) : (C, B) \in \overline{R}.$$

By $A \notin Ext(S)$ and Definition 7 we must have $A \notin S$ and some $A' \in I(A)$ such that $A' \notin S$. Consider arbitrary $B \in \mathsf{A}$ such that $(B, A') \in R$. Then by Definition 5 we get $(B, A) \in \overline{R}$. Since $Ext(S)$ defends A there must be some $C \in Ext(S)$ such that $(C, B) \in \overline{R}$. But then there is also $C' \in I(C)$ such that $(C', B) \in R$. From Definition 7 it follows that $C' \in S$, so that S defends A'. Since S is complete and B was arbitrary this implies $A' \in S$, contrary to assumption.
2) Assume $S \in \mathsf{Com}(\overline{\mathsf{A}}, \overline{R})$. We have to show $Res(S) \in \mathsf{Com}(\mathsf{A}, R)$. By Theorem 1 we know that $Res(S) \in \mathsf{Adm}(\mathsf{A}, R)$ so we only have to show completeness. Assume towards contradiction that there is some $A \in \mathsf{A} \setminus Res(S)$ such that:

$$\forall (B, A) \in (\mathsf{A}, R) : \exists C \in Res(S) : (C, B) \in (\mathsf{A}, R)$$

By Definition 7 we must have $A \notin S$, since otherwise $I(A) = \{A\}$ forcing $A \in Res(S)$. Consider arbitrary $B \in \overline{\mathsf{A}}$ such that $(B, A) \in R$. By Definition 5 there must be

$B' \in I(B), A' \in I(A)$ such that $(B', A') \in R$. Furthermore, since $A \in \mathsf{A}$, we have $I(A) = \{A\}$ so $A' = A$. It follows that $(B', A) \in R$. By the assumption that $Res(S)$ defends A we have $C \in Res(S)$ such that $(C, B') \in R$. Since $C \in Res(S)$, there must be some $C' \in S$ such that $C \in I(C')$. But then by Definition 5 we have $(C', B') \in \overline{R}$. Since B was arbitrary and $C' \in S$, S defends A in $(\overline{\mathsf{A}}, \overline{R})$. Hence, by completeness of S we get $A \in S$, contrary to assumption.

$\varepsilon = \mathsf{Pref}$: 1) Assume $S \in \mathsf{Pref}(\mathsf{A}, R)$. We have to show that $Ext(S) \in \mathsf{Pref}(\overline{\mathsf{A}}, \overline{R})$. Assume towards contradiction that there is some admissible $S' \supset Ext(S)$. Then by Definition 7 there is some $A \in S'$ with $I(A) \not\subseteq S$. Since $S \subseteq \mathsf{A}$ it follows from Definition 7 that $S \subseteq Ext(S)$ so that $Res(Ext(S)) = S$. Hence, $Res(S') \supset S$ (the inclusion is strict since $S \not\supseteq I(A) \subseteq Res(S')$), which by Theorem 1 contradicts the fact that $S \in \mathsf{Pref}(\mathsf{A}, R)$.
2) Assume $S \in \mathsf{Pref}(\overline{\mathsf{A}}, \overline{R})$. We have to show that $Res(S) \in \mathsf{Pref}(\mathsf{A}, R)$. Assume towards contradiction that there is some admissible $S' \supset Res(S)$. By Definition 7 we have $Ext(Res(S)) \supseteq S$. Hence, $Ext(S') \supseteq S$. Furthermore, since $S' \supset Res(S)$ there must be $A \in S' \setminus Res(S)$. Hence, from Definition 7 we get $Ext(S') \supset S$. But by Theorem 1 we have that $Ext(S') \in \mathsf{Adm}(\overline{\mathsf{A}}, \overline{R})$, contradicting $S \in \mathsf{Pref}(\overline{\mathsf{A}}, \overline{R})$.

$\varepsilon = \mathsf{Ground}$: 1) Assume $S \in \mathsf{Ground}(\mathsf{A}, R)$. We have to show that $Ext(S) \in \mathsf{Ground}(\overline{\mathsf{A}}, \overline{R})$. Assume towards contradiction that there is some complete $S' \subset Ext(S)$. Consider $Res(S')$. We have $Res(Ext(S)) = S$, so $Res(S') \subseteq S$. Since $Res(S')$ is complete, we must have $Res(S') = S$. Now, consider $A \in Ext(S) \setminus S'$. Consider arbitrary $B \in \overline{\mathsf{A}}$ such that $(B, A) \in \overline{R}$. By Definition 7 and $A \in Ext(S)$, we have $I(A) \subseteq S$. By Definition 5, we have some $A' \in I(A), B' \in I(B)$ such that $(A', B') \in R$. But since S is admissible and $I(A) \subseteq S$, this implies that there is $C \in S$ such that $(C, B') \in R$. However, from $I(C) = \{C\}$ and $Res(S') = S$ it follows that $C \in S'$. Hence, S' defends A against the attack from B. Since B was arbitrary and S' is complete, we get $A \in S'$, contrary to assumption.
2) Assume $S \in \mathsf{Ground}(\overline{\mathsf{A}}, \overline{R})$. By Definition 7, we see that $Ext(Res(S) \supseteq S$. Note that $Ext(Res(S)) \subseteq S$. To see this, let $A \in Ext(Res(S)$ be arbitrary. By Definition 7 we have $I(A) \subseteq Res(S)$. Hence, for every $A' \in I(A)$ there is some $A'' \in S$ such that $A' \in I(A'')$. Consider arbitrary $B \in \overline{\mathsf{A}}$ such that $(B, A) \in \overline{R}$. Then by Definition 5 there is some $A' \in I(A), B' \in I(B)$ such that $(B', A') \in R$. But then there is also $A'' \in S$ such that $(B, A'') \in \overline{R}$, which in turn means there is $C \in S$ such that $(C, B) \in \overline{R}$. Hence, S defends A, so that $A \in S$ follows from the fact that S is complete. □

6 Conclusion

In this paper, we have proposed a formalisation of enthymematic arguments in a framework of structured argumentation with Dung-style semantics. Representing arguments as sequences made up of formulas and rules, we considered sequences that for some reason do not satisfy those properties "regular" arguments are expected to satisfy. If a sequence can be extended to one or more regular arguments, we took it to be an enthymematic argument. We showed that this fits well with the traditional view on such arguments, whereby they are conceived of as arguments that are missing certain components (typically premises). In addition, we showed how this formalisation allows us to provide a semantics for enthymematic arguments in terms of abstract argumentation frameworks. Furthermore, we characterised the class of acceptable argument sets under admissible, complete, preferred, and grounded semantics, by relating acceptability over the regular argument with acceptability over the set of all sequences of rules and formulas (including incoherent and superfluous arguments alongside the regular and enthymematic ones).

The basic intuition behind our semantics is that an enthymematic argument corresponds to a set of regular arguments, namely the set of arguments corresponding to ways in which to "fill the gaps". This lead to the notion of a crater, which is based on a similar notion from the informal logic literature. In our opinion, the idea that enthymematic arguments have craters is very natural. Moreover, it leads naturally to a semantics where you attack an enthymematic argument if you attack at least one of the regular arguments in its crater. It seems to us that the resulting formalism has significant potential when it comes to integrating enthymematic arguments into the computational theory of argumentation. In future work, we plan to explore the definitions we have presented here in further depth, including an investigation of what happens when we add preferences to the formalism, in the style of ASPIC$^+$.

References

[Gilbert, 1991] Michael A. Gilbert. The enthymeme buster: A heuristic procedure for position exploration in dialogic dispute. *Informal Logic*, 13(3), 1991.

[Goddu, 2016] G. C. Goddu. On the very concept of an enthymeme. 2016.

[Modgil and Prakken, 2014] Sanjay Modgil and Henry Prakken. The *ASPIC*$^+$ framework for structured argumentation: a tutorial. *Argument & Computation*, 5(1):31–62, 2014.

[Paglieri and Woods, 2011] Fabio Paglieri and John Woods. Enthymemes: From reconstruction to understanding. *Argumentation*, 25(2):127–139, May 2011.

[Rahwan and Simari, 2009] Iyad Rahwan and Guillermo Simari, editors. *Argumentation in artificial intelligence*. Springer, 2009.

[Toulmin, 2003] Stephen Toulmin. *The Uses of Argument*. Cambridge University Press, 2 edition, 2003. First edition from 1958.

[Walton, 2008] Douglas Walton. The three bases for the enthymeme: A dialogical theory. *J. Applied Logic*, 6(3):361–379, 2008.