Enthymemes in Abstract Argumentation: Possible Worlds as Observers

Francesco Santini

Dept. of Mathematics and Computer Science, University of Perugia

AI^3 2026, Perugia, October 6–9, 2026



A.D. 1308
unipg
UNIVERSITÀ DEGLI STUDI
DI PERUGIA

Enthymemes

- An enthymeme is an argument with missing premises or a missing conclusion (Walton and Macagno, 2009).
- Aristotle's argument from sign: *"If the sky is clouded over, it is likely to rain."*
- Implicit premise: cloudy days are for the most part rainy.

The audience

An enthymeme is addressed to an *audience*: the *observers* that follow the debate without taking part in it. An enthymeme becomes an argument only for an observer that completes the missing part.

The framework changes from observer to observer.

Where does the missing part come from?

Dialogical context

What the debate has already made explicit.

Assumption: every observer supplies the same part (a simplification).

Common knowledge

The background that the speaker assumes to be shared.

Observers with different knowledge supply different parts, or none.

- Speaker and audience need not hold the same common knowledge (Mailly, 2016).
- This work addresses common knowledge, the only source that varies across the audience.
- Whether an enthymeme and its attacks belong to the framework of an observer depends on whether and how that observer completes the enthymeme.

Contributions

1. **AFEs**: Abstract Argumentation Frameworks with Enthymemes, which separate arguments from enthymemes and explicit attacks from enthymematic attacks.
2. **PAFEs**: a constellations semantics for AFEs, in which a possible world is the framework shared by a group of observers and the random variables are the completions of the enthymemes.
3. **Knowledge-Based PAFEs**: possible worlds indexed by the common knowledge of the audience: an observer that holds the knowledge needed by two enthymemes completes both, and an observer that lacks that knowledge completes neither.

Background: AFs and the constellations approach

- An **AF** (Dung, 1995) is a pair $\mathcal{F} = (\mathrm{A}, \mathrm{R})$ with $\mathrm{R} \subseteq \mathrm{A} \times \mathrm{A}$.
- A semantics σ maps each AF to a set $\sigma(\mathcal{F})$ of *extensions*. The examples use the stable semantics: $S \in \mathbf{st}(\mathcal{F})$ iff S is conflict-free and attacks every argument in $\mathrm{A} \setminus S$.

Probabilistic AF, constellations approach (Li, Oren and Norman, 2011)

- The uncertainty resides in the topology of the framework.
- Arguments and attacks carry a probability of existence, as independent events.
- Each induced AF has the probability of the corresponding selection of elements.
- $P_\sigma(B, \mathcal{F}_p)$ sums the probabilities of the induced AFs in which B is a σ-extension.

Why constellations: observers change the topology of the framework by completing enthymemes in different ways.

Abstract Argumentation Frameworks with Enthymemes

Definition (AFE)

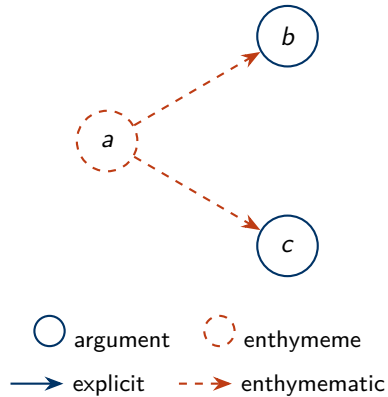
An AFE is a tuple $\mathcal{F}_e = (\mathsf{A}, \mathsf{E}, \mathsf{R}, \mathsf{R}_e)$ where

- A is the set of *arguments*, which are fully expressed;
- E is the set of *enthymemes*, with $\mathsf{A} \cap \mathsf{E} = \emptyset$; the *nodes* are $\mathsf{A} \cup \mathsf{E}$;
- R is the *explicit* attack relation: an explicit attack involves only parts that are explicit in both endpoints;
- R_e is the *enthymematic* attack relation: an enthymematic attack originates from or is directed to a missing part, so at least one endpoint is in E.

- If $\mathsf{E} = \emptyset$, an AFE is a Dung framework.
- An attack in R may involve an enthymeme: the explicit conclusion of an enthymeme may rebut the conclusion of another node.

Example: Bob

- *a*: "Bob is greedy", with the major premise and the conclusion both missing.
- Completion 1: "greedy people are unhappy, hence Bob is unhappy", in conflict with *b*: "Bob is joyful".
- Completion 2: "greedy people never share, hence Bob never shares anything", in conflict with *c*: "Bob lent me his car".
- Both conflicts belong to R_e.



- Only the enthymematic part of the graph, E and R_e, can be completed differently by different observers.

Completions and Probabilistic AFEs

- A framework records conflicts, not propositions: a completion of e is identified with the set of enthymematic attacks that the completion produces, a subset of $\mathrm{R}_e(e) = \{(x, y) \in \mathrm{R}_e \mid x = e \text{ or } y = e\}$.

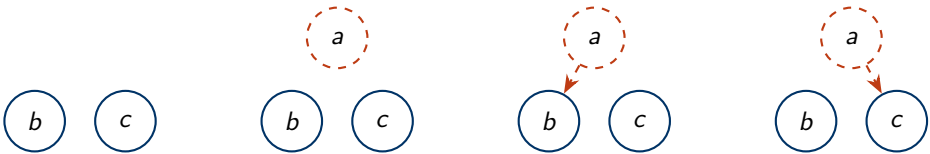
Definition (PAFE)

A PAFE $\mathcal{F}_{pe} = (\mathrm{A}, \mathrm{E}, \mathrm{R}, \mathrm{R}_e, P)$ assigns to each $e \in \mathrm{E}$ a distribution P_e over $2^{\mathrm{R}_e(e)} \cup \{\bot\}$. The completions C_e of e are the sets U with $P_e(U) > 0$.

- The observers are the elementary events: a possible world is the framework of an observer, shared by all the observers that end up with it.
- $P_e(U)$ is the fraction of observers whose completion of e produces exactly the conflicts in U; $P_e(\bot)$ is the fraction that leaves e incomplete.
- Node presence follows: e is absent for the observers with $\bot$, and present, with exactly the attacks in U, for those with a completion U, even $U = \emptyset$.
- Enthymematic attacks have no probability parameter of their own: their presence follows from the completions of their enthymematic endpoints.

Example: the completions of Bob's enthymeme

$\mathsf{A} = \{b, c\}$, $\mathsf{E} = \{a\}$, $\mathsf{R}_e(a) = \{(a, b), (a, c)\}$, and $\mathsf{C}_a = \{\emptyset, \{(a, b)\}, \{(a, c)\}\}$.



$\bot$: 0.4 $\emptyset$: 0.2 $\{(a, b)\}$: 0.2 $\{(a, c)\}$: 0.2

- With $\bot$, a is absent. With $\emptyset$, a is present without attacks.
- $P_a(\{(a, b), (a, c)\}) = 0$: in this example, the two completions of a exclude each other, so no observer produces both conflicts.
- A PAFE can exclude the pair, while a PrAF cannot: in a PrAF, a and both attacks occur together with probability $P_a \cdot P_{ab} \cdot P_{ac}$, which is positive as soon as each attack can occur.

Induced AFs

- A *profile* γ assigns to each $e \in \mathsf{E}$ a value $\gamma(e) \in \mathsf{C}_e \cup \{\bot\}$, with probability $\pi(\gamma) = \prod_{e \in \mathsf{E}} P_e(\gamma(e))$.
- The AF induced by γ contains
  - all the arguments and the enthymemes e with $\gamma(e) \neq \bot$;
  - the explicit attacks, and the enthymematic attacks r with $r \in \gamma(e)$ for every endpoint $e \in \mathsf{E}$ of r,

  restricted to the nodes present.
- $P_\sigma(B, \mathcal{F}_{pe}) = \sum_{\gamma \,:\, B \in \sigma(\mathcal{F}(\gamma))} \pi(\gamma)$ is the fraction of observers for which B is a σ-extension.

Proposition 1

The probabilities of the induced AFs sum to 1, and every induced AF has the same attacks as R between arguments of A.

Example: two enthymemes, independent

$\mathsf{A} = \{c\}$, $\mathsf{E} = \{a, b\}$, $\mathsf{R}_e = \{(a, c), (b, c)\}$, $P_a(\{(a, c)\}) = P_b(\{(b, c)\}) = 0.6$, $P_a(\bot) = P_b(\bot) = 0.4$.



$\mathcal{F}_1$: 0.16 $\mathcal{F}_2$: 0.24 $\mathcal{F}_3$: 0.24 $\mathcal{F}_4$: 0.36

- $\{c\}$ is stable in $\mathcal{F}_1$ only: $P_{\mathbf{st}}(\{c\}, \mathcal{F}_{pe}) = 0.16$, the fraction of observers that complete neither enthymeme.
- $\{a\}$ is stable in $\mathcal{F}_2$ only: $P_{\mathbf{st}}(\{a\}, \mathcal{F}_{pe}) = 0.24$.

Enthymemes are not independent

- The profiles of a PAFE complete distinct enthymemes independently.
- Suppose that a and b leave the same part implicit:
  - an observer that knows that part completes both enthymemes;
  - an observer that does not know that part completes neither.
- The worlds $\mathcal{F}_2$ and $\mathcal{F}_3$ should then have probability 0, not 0.24.
- Two enthymemes often leave the same part implicit: a speaker omits what the audience is supposed to know (Walton and Macagno, 2009).

Idea

Assign probability to the *knowledge items* that completing an enthymeme requires, and let the items decide which enthymemes are completed and which attacks hold.

Knowledge-Based PAFEs

Definition (Knowledge-Based PAFE)

$\mathcal{F}_k = (\mathsf{A}, \mathsf{E}, \mathsf{R}, \mathsf{R}_e, \mathsf{K}, P_{\mathsf{K}}, \kappa)$, where

- K is a finite set of *knowledge items*, each held by an observer independently of the others;
- $P_{\mathsf{K}} : \mathsf{K} \to (0, 1]$ gives, for each item k, the fraction $P_{\mathsf{K}}(k)$ of observers that hold k;
- $\kappa : \mathsf{E} \cup \mathsf{R}_e \to 2^{\mathsf{K}}$ specifies the items needed to complete each enthymeme and to produce each enthymematic attack.

- A possible world is a set $\mathsf{K}' \subseteq \mathsf{K}$ of items, shared by the observers that hold exactly those items, with probability $\pi(\mathsf{K}') = \prod_{k \in \mathsf{K}'} P_{\mathsf{K}}(k) \prod_{k \in \mathsf{K} \setminus \mathsf{K}'} (1 - P_{\mathsf{K}}(k))$.
- $\mathcal{F}(\mathsf{K}')$ contains the enthymemes e with $\kappa(e) \subseteq \mathsf{K}'$ and the enthymematic attacks r with $\kappa(r) \subseteq \mathsf{K}'$ between nodes present.
- The modeler elicits the items instead of the completions (Hosseini, Modgil and Rodrigues, 2016).

Shared items create correlation

Proposition 2

If every item is *private*, i.e., needed by a single enthymeme or a single attack, and every enthymematic attack has exactly one endpoint in E, then a Knowledge-Based PAFE induces the same distribution over AFs as a PAFE.

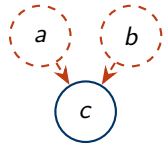
- Correlation therefore requires shared items.
- Neither model contains the other: only Knowledge-Based PAFEs can correlate distinct enthymemes (see slide 15), and only PAFEs can let the completions of one enthymeme exclude each other (see slide 16).

Example: two enthymemes, one shared item

Same AFE as before, with $\mathsf{K} = \{k\}$, $P_{\mathsf{K}}(k) = 0.6$ and $\kappa(x) = \{k\}$ for every $x \in \mathsf{E} \cup \mathsf{R}_e$.



$\mathcal{F}_1$: 0.4 (no k)



$\mathcal{F}_4$: 0.6 (holds k)

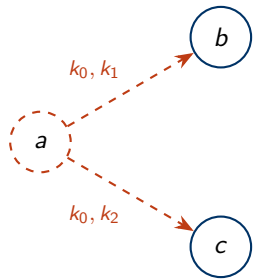
	independent (PAFE)	shared item
$P_{\mathbf{st}}(\{c\})$	0.16	0.4
$P_{\mathbf{st}}(\{a\})$	0.24	0

Ignoring the shared item underestimates $P_{\mathbf{st}}(\{c\})$ by a factor of 2.5 and assigns positive probability to $\{a\}$, which no observer accepts.

A limit: alternative reconstructions

Bob again: $\kappa(a) = \{k_0\}$, $\kappa((a,b)) = \{k_0, k_1\}$, $\kappa((a,c)) = \{k_0, k_2\}$, $P_K(k) = 0.5$ for every item.

- Half of the audience lacks k_0, so a and the attacks of a are absent.
- The other half forms four groups of 0.125: neither conflict, b only, c only, and both conflicts, for the observers with k_0, k_1 and k_2.
- If the two reconstructions of a are alternatives, as in slide 9, the group with both conflicts should not exist, and no choice of the items removes it.



Why

Acquiring an item can only add nodes and attacks, and the items are independent: two alternative reconstructions of the same missing part can never exclude each other.

Related work

- **Enthymemes in the constellations approach**: probabilities on attacks motivated by enthymematic arguments (Hunter, 2014).
- **Enthymemes and revision**: mismatch between the common knowledge of speaker and receiver (Mailly, 2016).
- **Independence questioned**: Li, Oren and Norman (2013).
- **Closest**: the likelihood of building an argument derives from those of holding the required beliefs (Hosseini et al., 2016).
- **Audiences as orderings of values** (Bench-Capon et al., 2007), not as sets of observers.

Conclusions and future work

Summary

- AFEs bring enthymemes into abstract argumentation.
- Two constellations semantics: possible worlds as the frameworks of groups of observers, then possible worlds indexed by the common knowledge of the audience.

Future work

- Conditions with negated items, where acquiring an item can remove an attack or a node, to let alternative reconstructions exclude each other and to represent *fallacies*, which an observer that spots the flaw rejects.
- The complexity of computing P_σ.

Thank you

Francesco Santini
francesco.santini@unipg.it