Preference-based argumentation systems

Leila Amgoud and Srdjan Vesić

IRIT - CNRS CRIL - CNRS

Preferences in argumentation frameworks

- An argument may be stronger than another:
    - is built from more certain information
    - refers to important goals
    - promotes more important value
    - . . .

Preferences in argumentation frameworks

- An argument may be stronger than another:
  - is built from more certain information
  - refers to important goals
  - promotes more important value
  - ...
- Need to take into account the strengths of arguments (captured by a preference relation $\geq \subseteq \mathcal{A} \times \mathcal{A}$)

Overview

1. Two roles of preferences in argumentation
   - Handling critical attacks
   - Refining the result
2. A framework integrating the two roles
3. Links with non-argumentative approaches

Two roles of preferences

Handling critical attacks

$$a \longrightarrow b$$

$$b > a$$

Two roles of preferences

Handling critical attacks

$b > a$

- We should accept b and reject a

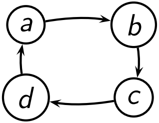
Two roles of preferences

Handling critical attacks

$b > a$

- We should accept b and reject a

Refining the result



$a > b$, $c > d$

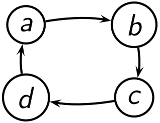
Two roles of preferences

Handling critical attacks

$$b > a$$

- We should accept b and reject a

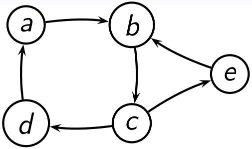
Refining the result



$$a > b, \ c > d$$

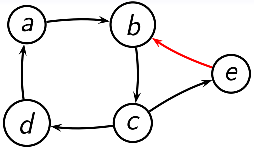
- Two stable extensions: $\{a, c\}$ and $\{b, d\}$
- However, $\{a, c\} \succ \{b, d\}$

The two roles are independent



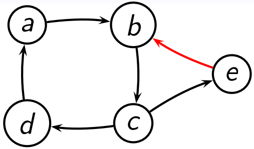
$a > b,\ c > d,\ b > e$

The two roles are independent



$a > b, \; c > d, \; b > e$

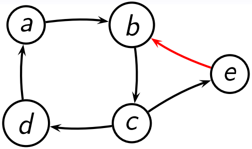
The two roles are independent



$$a > b,\ c > d,\ b > e$$

Argument e should be rejected

The two roles are independent

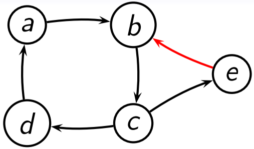


$$a > b,\ c > d,\ b > e$$

Argument e should be rejected

- The first role saves b from $e \Longrightarrow$ extensions $\{a, c\}$ and $\{b, d\}$

The two roles are independent



$$a > b,\ c > d,\ b > e$$

Argument e should be rejected

- The first role saves b from $e \Longrightarrow$ extensions $\{a, c\}$ and $\{b, d\}$
- The second role allows to refine: $\{a, c\} \succ \{b, d\}$

Rich PAF: integrating both roles of preferences

Input: $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq, \succeq)$

where $\geq \subseteq \mathcal{A} \times \mathcal{A}$ and $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$

Rich PAF: integrating both roles of preferences

Input: $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq, \succeq)$

$$\text{where } \geq \subseteq \mathcal{A} \times \mathcal{A} \text{ and } \succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$$

- First step: to handle critical attacks in the PAF $(\mathcal{A}, \mathcal{R}, \geq)$
    - To compute extensions $\mathcal{E}_1, \ldots, \mathcal{E}_n$

Rich PAF: integrating both roles of preferences

Input: $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq, \succeq)$

where $\geq \subseteq \mathcal{A} \times \mathcal{A}$ and $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$

- First step: to handle critical attacks in the PAF $(\mathcal{A}, \mathcal{R}, \geq)$
  - To compute extensions $\mathcal{E}_1, \ldots, \mathcal{E}_n$
- Second step: to use the relation $\succeq$ to compare the extensions
  - Example: $\mathcal{E} \succeq_d \mathcal{E}'$ iff $\forall x' \in \mathcal{E}' \setminus \mathcal{E}, \exists x \in \mathcal{E} \setminus \mathcal{E}'$ s.t. $x > x'$
  - $\mathcal{E}_i$ is an extension of $\mathcal{T}$ iff $\nexists \mathcal{E}_j$ s.t. $\mathcal{E}_j \succ \mathcal{E}_i$

Need for conflict-free extensions

Existing approaches may lead to non conflict-free extensions

$a\mathcal{R}b \quad b > a$

Need for conflict-free extensions

Existing approaches may lead to non conflict-free extensions

$a\mathcal{R}b \quad b > a$

Need for conflict-free extensions

Existing approaches may lead to non conflict-free extensions

$a\mathcal{R}b \quad b > a$

Need for conflict-free extensions

Existing approaches may lead to non conflict-free extensions

a b

$a\mathcal{R}b$ $b > a$

- $\mathcal{E} = \{a, b\}$

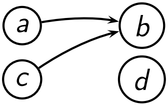
Need for conflict-free extensions

Existing approaches may lead to non conflict-free extensions

$a\mathcal{R}b \quad b > a$

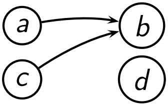
- $\mathcal{E} = \{a, b\}$
- An extension containing conflicting arguments

Need for comparing sets of arguments



$$a \approx c \qquad a > b$$
$$c > b \qquad d > b$$

Need for comparing sets of arguments

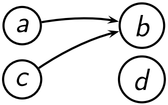


$a \approx c \quad a > b$

$c > b \quad d > b$

- Stable / preferred / grounded extension: $\{a, c, d\}$

Need for comparing sets of arguments



$$a \approx c \quad a > b$$
$$c > b \quad d > b$$

- Stable / preferred / grounded extension: $\{a, c, d\}$

- It is impossible to conclude that:
  - $\{a, c\} \succ \{b\}$
  - $\{d\} \succ \{b\}$
  - ...

A new approach for preference-based argumentation

Idea: to define new acceptability semantics that:

A new approach for preference-based argumentation

Idea: to define new acceptability semantics that:

- are based on preferences and attacks between arguments

A new approach for preference-based argumentation

Idea: to define new acceptability semantics that:

- are based on preferences and attacks between arguments
- generalize Dung's semantics

A new approach for preference-based argumentation

Idea: to define new acceptability semantics that:

- are based on preferences and attacks between arguments
- generalize Dung's semantics
- ensure conflict-free extensions

A new approach for preference-based argumentation

Idea: to define new acceptability semantics that:

- are based on preferences and attacks between arguments
- generalize Dung's semantics
- ensure conflict-free extensions
- allow to compare any pair of subsets of arguments

A new approach for preference-based argumentation

Definition (New semantics)

Let $(\mathcal{A}, \mathcal{R}, \geq)$ be a PAF. A semantics is defined by a dominance relation $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$.

The extensions of $(\mathcal{A}, \mathcal{R}, \geq)$ are the maximal elements of $\succeq$.

Definition (Maximal element)

$\mathcal{E} \in \mathcal{P}(\mathcal{A})$ is a maximal element of a dominance relation $\succeq$ iff $\forall \mathcal{E}' \in \mathcal{P}(\mathcal{A})$, $\mathcal{E} \succeq \mathcal{E}'$.

$\succeq_{max}$ = the set of all maximal elements wrt $\succeq$.

Generalizing stable semantics with preferences

Definition (Pref-stable semantics)

Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$. $\mathcal{E} \succeq_{st} \mathcal{E}'$ iff:

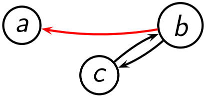
- $\mathcal{E}$ is conflict-free and $\mathcal{E}'$ is not conflict-free, or
- $\mathcal{E}$ and $\mathcal{E}'$ are conflict-free and $\forall a' \in \mathcal{E}' \setminus \mathcal{E}$, $\exists a \in \mathcal{E} \setminus \mathcal{E}'$ s.t. $(a\mathcal{R}a'$ and $a' \not> a)$ or $(a > a')$

Generalizing stable semantics with preferences

Definition (Pref-stable semantics)

Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$. $\mathcal{E} \succeq_{st} \mathcal{E}'$ iff:

- $\mathcal{E}$ is conflict-free and $\mathcal{E}'$ is not conflict-free, or
- $\mathcal{E}$ and $\mathcal{E}'$ are conflict-free and $\forall a' \in \mathcal{E}' \setminus \mathcal{E}$, $\exists a \in \mathcal{E} \setminus \mathcal{E}'$ s.t. $(a\mathcal{R}a'$ and $a' \not> a)$ or $(a > a')$



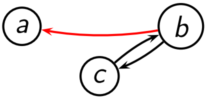
$a > b$

Generalizing stable semantics with preferences

Definition (Pref-stable semantics)

Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$. $\mathcal{E} \succeq_{st} \mathcal{E}'$ iff:

- $\mathcal{E}$ is conflict-free and $\mathcal{E}'$ is not conflict-free, or
- $\mathcal{E}$ and $\mathcal{E}'$ are conflict-free and $\forall a' \in \mathcal{E}' \setminus \mathcal{E}$, $\exists a \in \mathcal{E} \setminus \mathcal{E}'$ s.t. $(a \mathcal{R} a'$ and $a' \not> a)$ or $(a > a')$



$a > b$

- $\{a\} \succ_{st} \{b\}$
- $\emptyset \succ_{st} \{a, b, c\}$
- $\{b\} \succ_{st} \emptyset$
- $\ldots$
- $\succeq_{st,max} = \{\{a, c\}\}$

Generalizing stable semantics with preferences

Definition (Pref-stable semantics)

Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$. $\mathcal{E} \succeq_{st} \mathcal{E}'$ iff:

- $\mathcal{E}$ is conflict-free and $\mathcal{E}'$ is not conflict-free, or
- $\mathcal{E}$ and $\mathcal{E}'$ are conflict-free and $\forall a' \in \mathcal{E}' \setminus \mathcal{E}$, $\exists a \in \mathcal{E} \setminus \mathcal{E}'$ s.t. $(a\mathcal{R}a'$ and $a' \not> a)$ or $(a > a')$

Theorem

Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF.

- *The relation $\succeq_{st}$ generalizes stable semantics.*
- *For all $\mathcal{E} \in \succeq_{st,max}$, $\mathcal{E}$ is a maximal conflict-free subset of $\mathcal{A}$.*

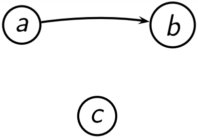
How to choose a dominance relation?

- Are there other relations that generalize this semantics?

How to choose a dominance relation?

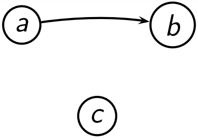
- Are there other relations that generalize this semantics?
- If yes,
  - what are the differences between them?
  - how to compare them?
  - are they all meaningful?

How to choose a dominance relation?

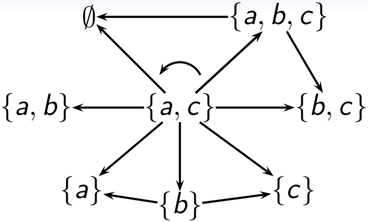


$c > b$

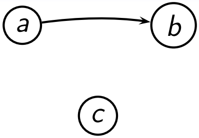
How to choose a dominance relation?



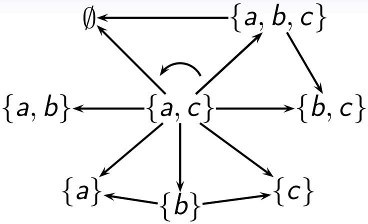
$c > b$



How to choose a dominance relation?

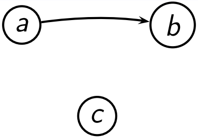


$c > b$

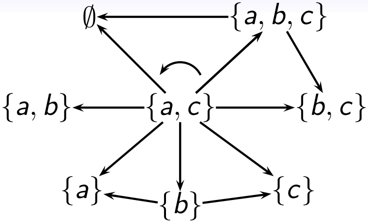


- This relation generalizes stable semantics

How to choose a dominance relation?

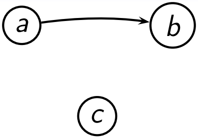


$c > b$

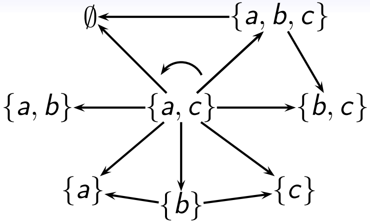


- This relation generalizes stable semantics
- However:
    - $\{a, b, c\} \succ \{b, c\}$

How to choose a dominance relation?

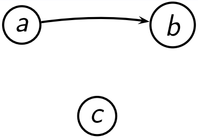


$c > b$

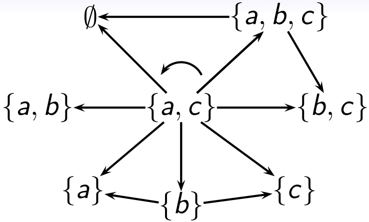


- This relation generalizes stable semantics
- However:
  - $\{a, b, c\} \succ \{b, c\}$
  - $\{b\} \succ \{a\}$

How to choose a dominance relation?



$c > b$



- This relation generalizes stable semantics
- However:
  - $\{a, b, c\} \succ \{b, c\}$
  - $\{b\} \succ \{a\}$
  - $\{b\} \succ \{c\}$

Postulates

Postulate (1)

$$\frac{\mathcal{E} \in \mathcal{CF} \qquad \mathcal{E}' \notin \mathcal{CF}}{\mathcal{E} \succ \mathcal{E}'}$$

Postulates

Postulate (1)

$$\frac{\mathcal{E} \in \mathcal{CF} \qquad \mathcal{E}' \notin \mathcal{CF}}{\mathcal{E} \succ \mathcal{E}'}$$

Postulate (2)

Let $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}$.

$$\frac{\mathcal{E} \succeq \mathcal{E}'}{\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}}$$

$$\frac{\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}}{\mathcal{E} \succeq \mathcal{E}'}$$

Postulates

Postulate (1)

$$\frac{\mathcal{E} \in \mathcal{CF} \qquad \mathcal{E}' \notin \mathcal{CF}}{\mathcal{E} \succ \mathcal{E}'}$$

Postulate (2)

Let $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}$.

$$\frac{\mathcal{E} \succeq \mathcal{E}'}{\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}}$$

$$\frac{\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}}{\mathcal{E} \succeq \mathcal{E}'}$$

Postulate (3)

Let $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}$ *and* $\mathcal{E} \cap \mathcal{E}' = \emptyset$.

$$\frac{(\exists a' \in \mathcal{E}')(\forall a \in \mathcal{E}) \quad \neg(a \mathcal{R} a' \wedge a' \not\succ a) \wedge a \not\succ a'}{\neg(\mathcal{E} \succeq \mathcal{E}')}$$

Postulates

Postulate (1)

$$\frac{\mathcal{E} \in \mathcal{CF} \qquad \mathcal{E}' \notin \mathcal{CF}}{\mathcal{E} \succ \mathcal{E}'}$$

Postulate (2)

Let $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}$.

$$\frac{\mathcal{E} \succeq \mathcal{E}'}{\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}}$$

$$\frac{\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}}{\mathcal{E} \succeq \mathcal{E}'}$$

Postulate (3)

Let $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}$ and $\mathcal{E} \cap \mathcal{E}' = \emptyset$.

$$\frac{\begin{array}{c}(\exists a' \in \mathcal{E}')(\forall a \in \mathcal{E}) \\ \neg(a\mathcal{R}a' \wedge a' \not\geq a) \wedge a \not\geq a'\end{array}}{\neg(\mathcal{E} \succeq \mathcal{E}')}$$

Postulate (4)

Let $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}$ and $\mathcal{E} \cap \mathcal{E}' = \emptyset$.

$$\frac{\begin{array}{c}(\forall a' \in \mathcal{E}')(\exists a \in \mathcal{E}) \\ (a\mathcal{R}a' \wedge a' \not\geq a) \text{ or } (a\mathcal{R}a' \wedge a > a')\end{array}}{\mathcal{E} \succeq \mathcal{E}'}$$

Properties of relations satisfying the postulates

Theorem

If $\succeq$ satisfies Postulates 1-4, then $\succeq$ generalizes stable semantics.

Theorem

If $\succeq$ and $\succeq'$ both satisfy Postulates 1-4, then $\succeq_{max} = \succeq'_{max}$.

...

Preferred sub-theories (Brewka'89)

Let $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ be a stratified propositional knowledge base.

Preferred sub-theories (Brewka'89)

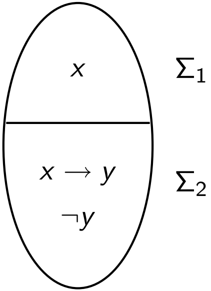
Let $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ be a stratified propositional knowledge base.

Definition

Let $\mathcal{S} \subseteq \Sigma$ and $\mathcal{S}_i = \mathcal{S} \cap \Sigma_i$. $\mathcal{S}$ is a preferred sub-theory iff for every $1 \leq k \leq n$, $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_k$ is a maximal consistent set in $\Sigma_1 \cup \ldots \cup \Sigma_k$

Preferred sub-theories (Brewka'89)

Let $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ be a stratified propositional knowledge base.



Definition

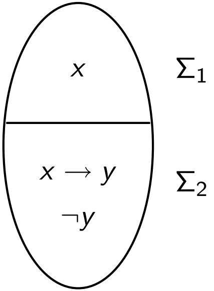
Let $\mathcal{S} \subseteq \Sigma$ and $\mathcal{S}_i = \mathcal{S} \cap \Sigma_i$.
$\mathcal{S}$ is a preferred sub-theory iff for every $1 \leq k \leq n$, $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_k$ is a maximal consistent set in $\Sigma_1 \cup \ldots \cup \Sigma_k$

Preferred sub-theories (Brewka'89)

Let $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ be a stratified propositional knowledge base.

Definition

Let $\mathcal{S} \subseteq \Sigma$ and $\mathcal{S}_i = \mathcal{S} \cap \Sigma_i$. $\mathcal{S}$ is a preferred sub-theory iff for every $1 \leq k \leq n$, $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_k$ is a maximal consistent set in $\Sigma_1 \cup \ldots \cup \Sigma_k$



$\mathcal{S}_1 = \{x, x \rightarrow y\}$
$\mathcal{S}_2 = \{x, \neg y\}$

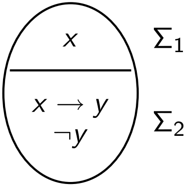
Preferred sub-theories and stable extensions of basic PAF

Let $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ be a stratified knowledge base

Theorem

There is a bijection between the set of preferred sub-theories of Σ and the set of stable extensions of $(\mathtt{Arg}(\Sigma), Undercut, \geq_{wlp})$.

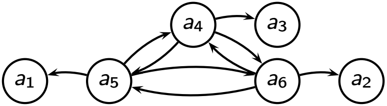
Preferred sub-theories and stable extensions of basic PAF



$a_1 : (\{x\}, x)$ $\quad$ $a_2 : (\{\neg y\}, \neg y)$

$a_3 : (\{x \rightarrow y\}, x \rightarrow y)$ $\quad$ $a_4 : (\{x, \neg y\}, x \wedge \neg y)$

$a_5 : (\{\neg y, x \rightarrow y\}, \neg x)$ $\quad$ $a_6 : (\{x, x \rightarrow y\}, y)$



$\mathcal{S}_1 = \{x, x \rightarrow y\}$

$\mathcal{S}_2 = \{x, \neg y\}$

$a_1 > a_2, a_3, a_4, a_5, a_6$

$\mathcal{E}_1 = \mathtt{Arg}(\mathcal{S}_1) = \{a_1, a_3, a_6, \ldots\}$

$\mathcal{E}_2 = \mathtt{Arg}(\mathcal{S}_2) = \{a_1, a_2, a_4, \ldots\}$

Democratic sub-theories and stable extensions of Rich PAF

- More general case: $\trianglerighteq \subseteq \Sigma \times \Sigma$ is not total

Definition (Cayrol & Royer & Saurel'93)

Given $(\Sigma, \trianglerighteq)$, a set $\mathcal{S} \subseteq \Sigma$ is a democratic sub-theory iff $\mathcal{S}$ is consistent and $(\nexists \mathcal{S}' \subseteq \Sigma)$ s.t. $\mathcal{S}'$ is consistent and $\mathcal{S}' \succeq_d \mathcal{S}$.

Democratic sub-theories and stable extensions of Rich PAF

- More general case: $\trianglerighteq \subseteq \Sigma \times \Sigma$ is not total

Definition (Cayrol & Royer & Saurel'93)

Given $(\Sigma, \trianglerighteq)$, a set $\mathcal{S} \subseteq \Sigma$ is a democratic sub-theory iff $\mathcal{S}$ is consistent and $(\nexists \mathcal{S}' \subseteq \Sigma)$ s.t. $\mathcal{S}'$ is consistent and $\mathcal{S}' \succeq_d \mathcal{S}$.

Theorem

There is a bijection between the set of democratic sub-theories of Σ and the set of stable extensions of the rich PAF $(\texttt{Arg}(\Sigma), \mathit{Undercut}, \geq_{gwlp}, \succeq_d)$.

Conclusion

- Clear distinction between two roles of preferences
- Rich model that takes into account both roles of preferences
- Novel approach for handling critical attacks
- Links with non-argumentative approaches for inconsistency handling

Conclusion

- Clear distinction between two roles of preferences
- Rich model that takes into account both roles of preferences
- Novel approach for handling critical attacks
- Links with non-argumentative approaches for inconsistency handling

Thank you