Artificial Intelligence 1

Prof. Dr. Frank Hopfgartner
Dr. Matthias Horbach

Institute for Web Science and Technologies (WeST)
University of Koblenz

Oberview

1. Introduction
2. Classical logics and Prolog
3. Search and automatic planning
4. Knowledge representation and reasoning
   - Default logic
   - Answer set programming
   - Argumentation
5. Agents and multi agent systems
6. Summary and conclusion

Rules and arguments 1/3

- Rules such as defaults and logical rules are the foundation of knowledge representation
- The central challenge is how to resolve conflicts between rules („Does the penguin fly, yes or no?“)
- In default logic and answer set programming, this is done by default assumptions
- Another approach can be obtained by considering *arguments* and *counterarguments*

Rules and arguments 2/3

- An *argument* is a set of rules (the *premise*) that entails some statement (the *conclusion*)
- For example

$$
\begin{aligned}
&bird(tweety) \leftarrow penguin(tweety)\\
&flies(tweety) \leftarrow bird(tweety), \texttt{not}\ \neg flies(tweety)\\
&penguin(tweety)
\end{aligned}
$$

  is an argument for the claim that Tweety flies
- ...and

$$
\begin{aligned}
&\neg flies(tweety) \leftarrow penguin(tweety)\\
&penguin(tweety)
\end{aligned}
$$

  is an argument for the claim that Tweety does not fly (a counterargument).

Rules and arguments 3/3

- The argument claiming that Tweety flies „needs“ the default assumption that Tweety does not not fly; it is therefore *weaker*
- The argument claiming that Tweety does not fly therefore *attacks* the other argument

Tweety flies because he is a bird and birds typically fly

Tweety does not fly because he is a penguin

Abstract argumentation frameworks

- Let's ignore the interior of arguments and focus on the relationship between arguments and counterarguments
- This can be represented as a directed graph
  - Arguments are nodes in the graph
  - A directed edge represented an attack of one argument on the other

Definition

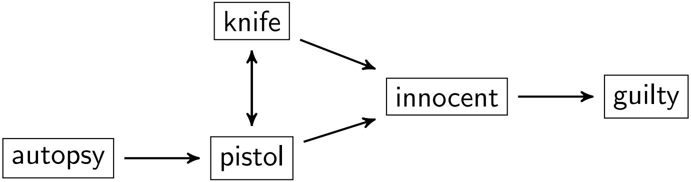
An *abstract argumentation framework* $\mathrm{AF} = (\mathrm{Arg}, \mathrm{R})$ is a graph with nodes Arg and edges $\mathrm{R} \subseteq \mathrm{Arg} \times \mathrm{Arg}$.

[Phan Ming Dung. On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games. Artificial Intelligence 77(2):321–358, 1995]

Example 1/5

John is accused of the murder of Mary. However, a person is assumed innocent unless proven guilty. There are two witness reports of Carl and Dave. Carl says that John killed Mary with a knife. Dave says that John killed Mary with a pistol. The autopsy of Mary confirms that she has no gunshot wound.

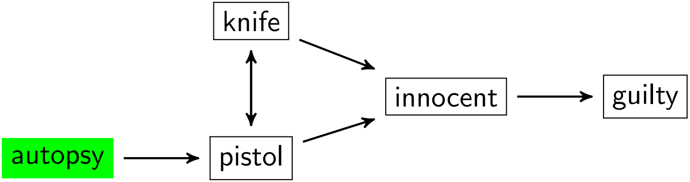
Formalisation as argumentation framework:



Question: Which arguments are acceptable?

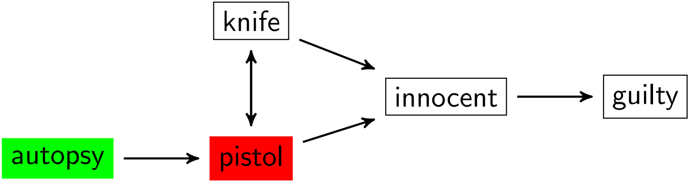
Example 2/5

1.) Arguments that are not attacked should be accepted

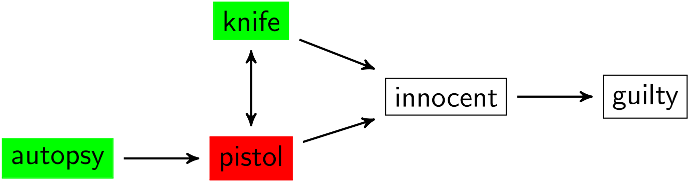


Example 2/5

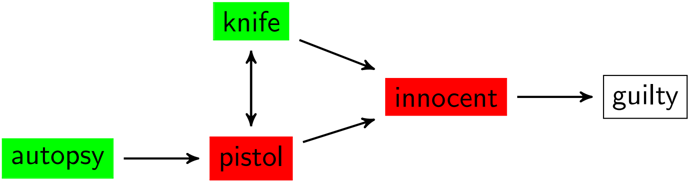
2.) Arguments that are attacked by an accepted argument, should not be accepted



3.) Arguments that are only attacked by unacceptable arguments, should be accepted

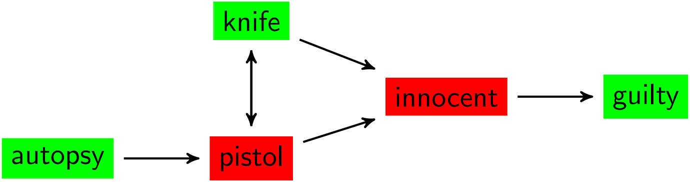


2.) Arguments that are attacked by an accepted argument, should not be accepted



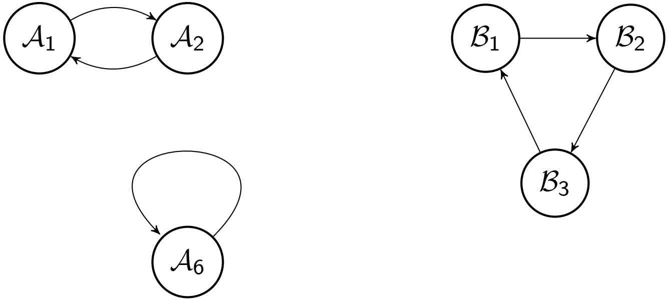
Example 5/5

3.) Arguments that are only attacked by unacceptable arguments, should be accepted



Problematic cases

What to do if the graphs are more complicated?



Conflict freeness, Defense

Let $\mathrm{AF} = (\mathrm{Arg}, \mathrm{R})$ be an argumentation framework.

Wanted: subset $E \subseteq \mathrm{Arg}$ of acceptable arguments.

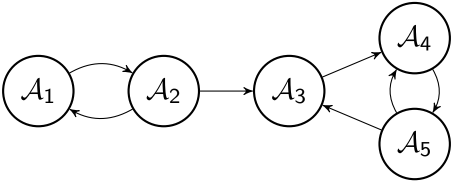
Definition

$E \subseteq \mathrm{Arg}$ is called *conflict-free* iff for all $\mathcal{A}, \mathcal{B} \in E$, $(\mathcal{A}, \mathcal{B}) \notin R$.

Definition

$E \subseteq \mathrm{Arg}$ *defends* $\mathcal{A} \in \mathrm{Arg}$, iff for all $\mathcal{B} \in \mathrm{Arg}$ with $(\mathcal{B}, \mathcal{A}) \in R$ there is a $\mathcal{C} \in E$ with $(\mathcal{C}, \mathcal{B}) \in R$.

Example

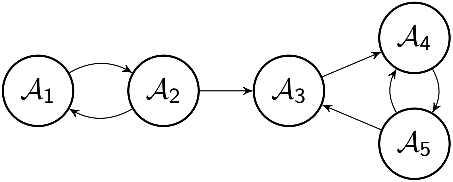


- $E_1 = \{\mathcal{A}_1, \mathcal{A}_3\}$ is conflict-free
- $E_2 = \{\mathcal{A}_3, \mathcal{A}_4, \mathcal{A}_5\}$ is not conflict-free
- $E_3 = \{\mathcal{A}_1, \mathcal{A}_4\}$ defends $\mathcal{A}_3$
- $E_4 = \{\mathcal{A}_5\}$ defends $\mathcal{A}_5$ (itself)

Admissibility 1/3

Definition

$E \subseteq \text{Arg}$ is *admissible* iff E is conflict-free and every $\mathcal{A} \in E$ is defended by E.



- $E_1 = \{\mathcal{A}_1, \mathcal{A}_3\}$ is not admissible
- $E_2 = \{\mathcal{A}_3, \mathcal{A}_4, \mathcal{A}_5\}$ is not admissible
- $E_3 = \{\mathcal{A}_1, \mathcal{A}_4\}$ is not admissible
- $E_4 = \{\mathcal{A}_5\}$ ist admissible
- $E_5 = \{\mathcal{A}_2, \mathcal{A}_4\}$ is admissible

Admissibility 2/3

Let $\mathrm{AF} = (\mathrm{Arg}, \mathrm{R})$ be an argumentation framework.

Theorem

If E is admissible and defends $\mathcal{A} \in \mathrm{Arg}$ then $E \cup \{\mathcal{A}\}$ is admissible.

Proof.

We have to show that $E_1 = E \cup \{\mathcal{A}\}$ is conflict-free and defends all its elements.

1. Conflict-freeness: Assume that E_1 is *not* conflict-free. As E is conflict-free, the conflict must be between E and $\mathcal{A}$. So there is a $\mathcal{B} \in E$ and either (i) $(\mathcal{B}, \mathcal{A}) \in R$ or (ii) $(\mathcal{A}, \mathcal{B}) \in R$:
   - (i) If $(\mathcal{B}, \mathcal{A}) \in R$ then (as E defends $\mathcal{A}$) there is a $\mathcal{C} \in E$ s.t. $(\mathcal{C}, \mathcal{B}) \in R$. This is impossible because E is conflict-free.
   - (ii) If $(\mathcal{A}, \mathcal{B}) \in R$ then (as E is admissible) E defends $\mathcal{B}$. So there is a $\mathcal{C} \in E$ s.t. $(\mathcal{C}, \mathcal{A}) \in R$, and (as E defends $\mathcal{A}$) a $\mathcal{D} \in E$ s.t. $(\mathcal{D}, \mathcal{C}) \in R$. But E is conflict-free, so this too is impossible.

   It follows that E_1 is conflict-free.

Admissibility 3/3

2. $E_1 = E \cup \{\mathcal{A}\}$ defends all its elements:
   As E was admissible, every $\mathcal{B} \in E$ is defended by E_1 as well.
   By assumption, $\mathcal{A}$ is defended by E (and therefore E_1). $\square$

Semantics for argumentation frameworks

- Question: are admissible sets of arguments a meaningful definition of "set of acceptable arguments"?
- Admissibility is very weak: the empty set $\emptyset$ is always admissible
- Argumentation *semantics* make this notion more precise
  - preferred semantics
  - complete semantics
  - grounded semantics
  - stable semantics

Preferred semantics

Let $\mathrm{AF} = (\mathrm{Arg}, \mathrm{R})$ be an argumentation framework.

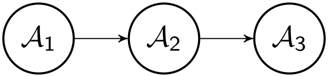
Definition

$E \subseteq \mathrm{Arg}$ is a *preferred extension* of AF iff

1. E is admissible and
2. there is no other admissible $E' \subseteq \mathrm{Arg}$ with $E \subsetneq E'$.

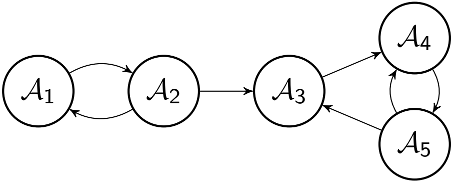
A preferred extension is therefore a maximal set of acceptable arguments.

Preferred semantics - Example 1



- Admissible sets: $\emptyset, \{\mathcal{A}_1\}, \{\mathcal{A}_1, \mathcal{A}_3\}$
- Preferred extension: $\{\mathcal{A}_1, \mathcal{A}_3\}$

Preferred semantics - Example 2



- Admissible sets:
  $\emptyset, \{\mathcal{A}_1\}, \{\mathcal{A}_2\}, \{\mathcal{A}_5\}, \{\mathcal{A}_2, \mathcal{A}_4\}, \{\mathcal{A}_2, \mathcal{A}_5\}, \{\mathcal{A}_1, \mathcal{A}_5\}$
- Preferred extensions: $\{\mathcal{A}_2, \mathcal{A}_4\}, \{\mathcal{A}_2, \mathcal{A}_5\}$, $\{\mathcal{A}_1, \mathcal{A}_5\}$

Existence of preferred extensions

Let $\mathrm{AF} = (\mathrm{Arg}, \mathrm{R})$ be an argumentation framework.

Theorem

AF *has at least one preferred extension.*

Proof.

Observe that $\emptyset$ is always admissible. Furthermore, the set of admissible sets $\mathcal{E} = \{E_1, \ldots, E_n\}$ of AF is always finite (if AF is finite). Therefore there is $j \in \{1, \ldots, n\}$ such that E_j has no strict superset in $\mathcal{E}$. Then E_j is a preferred extension. $\square$

Characteristic function

Let $\mathrm{AF} = (\mathrm{Arg}, \mathrm{R})$ be an argumentation framework.

Definition

The *characteristic function* F_{AF} of AF is the function $F_{\mathrm{AF}} : 2^{\mathrm{Arg}} \to 2^{\mathrm{Arg}}$ with

$$F_{\mathrm{AF}}(S) = \{\mathcal{A} \in \mathrm{Arg} \mid S \text{ defends } \mathcal{A}\}$$

Note: 2^X is the power set of X.

Characteristic function - Example

$$\mathcal{A}_1 \longrightarrow \mathcal{A}_2 \longrightarrow \mathcal{A}_3$$

- $F_{\mathsf{AF}}(\emptyset) = \{\mathcal{A}_1\}$
- $F_{\mathsf{AF}}(\{\mathcal{A}_1\}) = \{\mathcal{A}_1, \mathcal{A}_3\}$
- $F_{\mathsf{AF}}(\{\mathcal{A}_1, \mathcal{A}_3\}) = \{\mathcal{A}_1, \mathcal{A}_3\}$
- $F_{\mathsf{AF}}(\{\mathcal{A}_2\}) = \{\mathcal{A}_1\}$

Characteristic function and the grounded semantics

Theorem

Let E be conflict-free. Then E is admissible iff $E \subseteq F_{\mathsf{AF}}(E)$.

Theorem

F_{AF} is monotonous:

$$\emptyset \subseteq F_{\mathsf{AF}}(\emptyset) \subseteq F_{\mathsf{AF}}(F_{\mathsf{AF}}(\emptyset)) \subseteq \ldots$$

Definition

E is called the *grounded extension* of AF iff

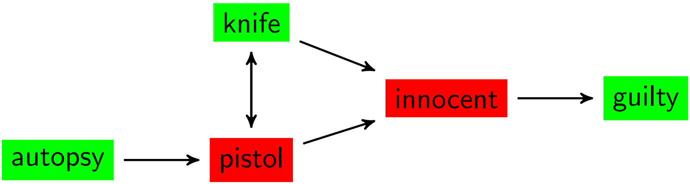
$$E = \bigcup_{i=1}^{\infty} F^{i}_{\mathsf{AF}}(\emptyset)$$

Grounded semantics - example 1/2

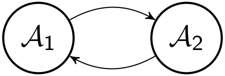
Computing the grounded extension corresponds exactly to the intuitive procedure from the beginning:

1. Arguments that are not attacked should be accepted ($F_{\mathsf{AF}}(\emptyset)$)
2. Arguments that are attacked by an accepted argument, should not be accepted
3. Arguments that are only attacked by unacceptable arguments, should be accepted (i. e., those that are defended by previously accepted arguments ($F^i_{\mathsf{AF}}(\emptyset)$))

Grounded semantics - example 2/2



$F_{\mathsf{AF}}(\emptyset) = \{\text{autopsy}\}$
$F_{\mathsf{AF}}(\{\text{autopsy}\}) = \{\text{autopsy}, \text{knife}\}$
$F_{\mathsf{AF}}(\{\text{autopsy}, \text{knife}\}) = \{\text{autopsy}, \text{knife}, \text{guilty}\}$
$F_{\mathsf{AF}}(\{\text{autopsy}, \text{knife}, \text{guilty}\}) = \{\text{autopsy}, \text{knife}, \text{guilty}\}$
...



$F_{\mathsf{AF}}(\emptyset) = \emptyset$

...

Grounded extension is empty

Grounded semantics - properties

Theorem

Every AF *has exactly one grounded extension (that could be empty).*

Proof.

We have already seen that F_{AF} is monotonous:

$$\emptyset \subseteq F_{\mathsf{AF}}(\emptyset) \subseteq F_{\mathsf{AF}}(F_{\mathsf{AF}}(\emptyset)) \subseteq \dots$$

As Arg is finite, necessarily we have $F^{i}_{\mathsf{AF}}(\emptyset) = F^{i+1}_{\mathsf{AF}}(\emptyset)$ for some i.
As the iterative application of the F_{AF} is a deterministic procedure, the grounded extension is unique

$$E = \bigcup_{i=1}^{\infty} F^{i}_{\mathsf{AF}}(\emptyset)$$

and exists. □

Complete semantics

Let $\mathrm{AF} = (\mathrm{Arg}, \mathrm{R})$ be an argumentation framework

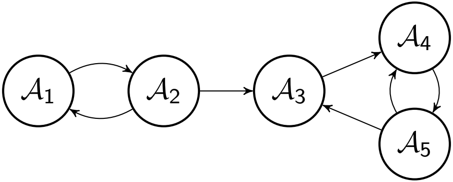
Definition
E is a *complete extension* iff $E = F_{\mathrm{AF}}(E)$.

Hence, a complete extension is a set of arguments that defends itself *and* contains all arguments it defends.

Theorem
Every AF *has at least one complete extension.*

Proof.
Every AF has a grounded extension, and if E is grounded then $F_{\mathrm{AF}}(E) = E$, so E is also complete. □

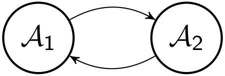
Complete semantics - example



$F_{\mathsf{AF}}(\emptyset) = \emptyset$
$F_{\mathsf{AF}}(\{\mathcal{A}_1\}) = \{\mathcal{A}_1\}$
$F_{\mathsf{AF}}(\{\mathcal{A}_2\}) = \{\mathcal{A}_2\}$
$F_{\mathsf{AF}}(\{\mathcal{A}_5\}) = \{\mathcal{A}_5\}$
$F_{\mathsf{AF}}(\{\mathcal{A}_1, \mathcal{A}_5\}) = \{\mathcal{A}_1, \mathcal{A}_5\}$
$F_{\mathsf{AF}}(\{\mathcal{A}_2, \mathcal{A}_5\}) = \{\mathcal{A}_2, \mathcal{A}_5\}$
$F_{\mathsf{AF}}(\{\mathcal{A}_2, \mathcal{A}_4\}) = \{\mathcal{A}_2, \mathcal{A}_4\}$

Complete extensions: $\emptyset$, $\{\mathcal{A}_1\}$, $\{\mathcal{A}_2\}$, $\{\mathcal{A}_5\}$, $\{\mathcal{A}_1, \mathcal{A}_5\}$, $\{\mathcal{A}_2, \mathcal{A}_5\}$, $\{\mathcal{A}_2, \mathcal{A}_4\}$

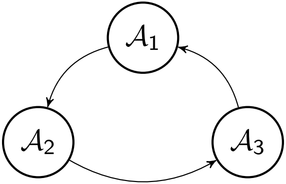
Complete semantics - Another example



$F_{\mathsf{AF}}(\emptyset) = \emptyset$
$F_{\mathsf{AF}}(\{\mathcal{A}_1\}) = \{\mathcal{A}_1\}$
$F_{\mathsf{AF}}(\{\mathcal{A}_2\}) = \{\mathcal{A}_2\}$

Complete extensions: $\emptyset$, $\{\mathcal{A}_1\}$, $\{\mathcal{A}_2\}$

Complete semantics - One more example



$F_{\mathsf{AF}}(\emptyset) = \emptyset$
$F_{\mathsf{AF}}(\{\mathcal{A}_1\}) = \{\mathcal{A}_3\}$
$F_{\mathsf{AF}}(\{\mathcal{A}_3\}) = \{\mathcal{A}_2\}$
$F_{\mathsf{AF}}(\{\mathcal{A}_2\}) = \{\mathcal{A}_1\}$

Complete extensions: $\emptyset$

Stable semantics

Let $\mathrm{AF} = (\mathrm{Arg}, \mathrm{R})$ be an argumentation framework.

Definition

E is a *stable extension* iff E is conflict-free and E attacks every argument in $\mathrm{Arg} \setminus E$.

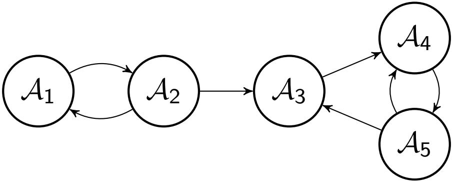
A stable extension is a very aggressive point of view: all non-acceptable arguments are explicitly attacked.

Hence

Theorem

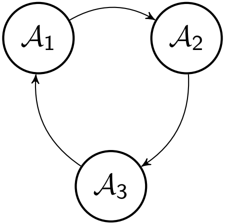
There are AF *without stable stable extensions.*

Stable semantics - example



stable extensions: $\{\mathcal{A}_2, \mathcal{A}_4\}, \{\mathcal{A}_2, \mathcal{A}_5\}$, $\{\mathcal{A}_1, \mathcal{A}_5\}$

Stable semantics - another example



- There are no stable extensions
  - Every singleton set does not attacked the other two arguments
  - Every two-element set is not conflict-free
- But note that $\emptyset$ is a complete extension (and also grounded and preferred)

Relationship between the semantics 1/2

Let $\mathsf{AF} = (\mathsf{Arg}, \mathsf{R})$ be an argumentation framework

Theorem

If E is stable then E is preferred.

Proof.

Let E be a stable extension. We have to show that E is admissible and there is no larger admissible set.

- *Admissibility*: E is conflict-free by definition. As E attacks all arguments outside of E it automatically defends all elements of E.
- *Maximality*: as E attacks all arguments outside of E adding another argument to E would raise a conflict.

□

Relationship between the semantics 2/2

Let $\mathsf{AF} = (\mathsf{Arg}, \mathsf{R})$ be an argumentation framework.

Theorem

If E is preferred then E is complete.

Theorem

If E is grounded then E is complete.

$$\begin{tikzcd}
\text{Stable semantics} \arrow[dr] & & & \\
& \text{Preferred semantics} \arrow[dr] & & \text{Grounded semantics} \arrow[dl] \\
& & \text{Complete semantics} &
\end{tikzcd}$$

Outlook

- There are many more semantics besides the one we discussed
  - Semi-stable semantics
  - Ideal semantics
  - Stage semantics
  - CF2 semantics
  - ...
- Computing an extension is usually a complex operation
  - Example: "Is E a preferred extension" is coNP-complete
  - Evaluation of implementations:
    `http://argumentationcompetition.org`
- Abstract argumentation subsumes many other approaches to plausible reasoning (such as default logic and answer set programming)

Chapter 4.3: Argumentation

Summary

Chapter 4.3: Summary

- Abstract argumentation frameworks $\mathsf{AF} = (\mathsf{Arg}, \mathsf{R})$
- Conflict-freeness, defense, admissibility
- Characteristic function
- Semantics
  - Preferred semantics
  - Grounded semantics
  - Complete semantics
  - Stable semantics