Admissibility in the Abstract Dialectical Framework

Polberg Sylwia, Johannes Peter Wallner and Stefan Woltran

September 16, 2013
CLIMA XIV

Abstract Argumentation

Argumentation

...the study of processes “concerned with how assertions are proposed, discussed, and resolved in the context of issues upon which several diverging opinions may be held”.
[Bench–Capon and Dunne, Argumentation in AI, AIJ 171:619–641, 2007]

Approaches

- Logic–based: arguments have an internal structure
- **Abstract**: arguments are abstract, we focus on relations

What is it all about?

Main ingredients

- Framework: represents arguments and relations between them
- Semantics: requirements and methods for choosing acceptable arguments (extensions) or labelings

Frameworks

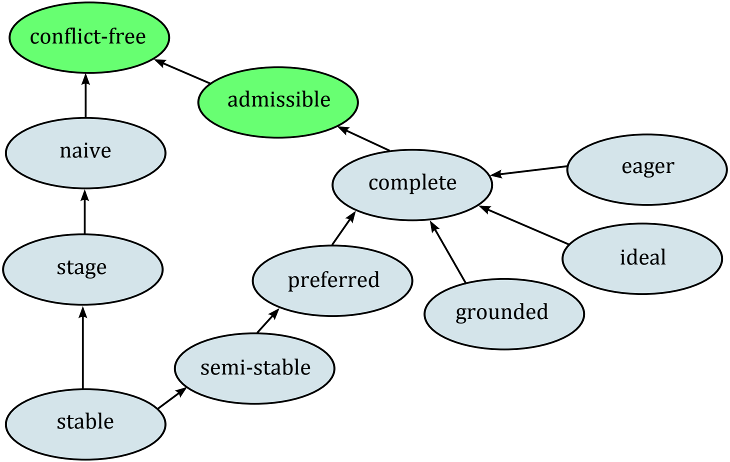
- Dung Framework AF
- AF generalizations: bipolar, recursive, weighted...
- ...and Abstract Dialectical Framework ADF

Semantics

Semantics grasp what we consider "rational", for example:

- Chosen arguments cannot be conflicting one with another or need form an opinion we can defend
- We maximize the amount of arguments we can accept or disprove
- We reject or accept circular reasoning

Relations between semantics



How to read: $\sigma \rightarrow \tau$ means that extensions of σ semantics are also extensions of τ semantics.

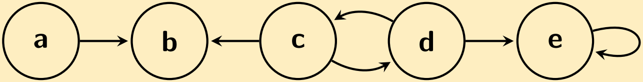
Roadmap

Dung framework

Dung framework

A **Dung abstract argumentation framework**, or a **Dung Framework** is a pair (A, R), where A is a set of arguments and $R \subseteq A \times A$ represents the attack relation.

Example



Let $AF = (A, R)$ be a Dung framework.

Conflict–freeness

Let (A, R) be a Dung framework. A set $S \subseteq A$ is **conflict–free** in AF iff there are no $a, b \in S$ s.t. $(a, b) \in R$.

Different flavours of admissibility: Dung I

Defense

An argument $a \in A$ is **defended** by a set $S \subseteq A$ in AF, if for each $b \in A$ s.t. $(b, a) \in R$, there exists $c \in S$ s.t. $(c, b) \in R$.

Standard extension–based definition

A conflict–free extension S is an **admissible extension** of AF if each $a \in S$ is defended by S in AF.

Characteristic function definition

The **characteristic function** of a Dung framework AF $F_{AF} : 2^A \rightarrow 2^A$ is defined as follows:

$$F_{AF}(S) = \{a \mid a \text{ is defended by } S \textit{ in } AF\}$$

A set $S \subseteq A$ is an **admissible extension** of AF iff it is conflict–free and $S \subseteq F_{AF}(S)$.

Different flavours of admissibility: Dung II

Legal labeling

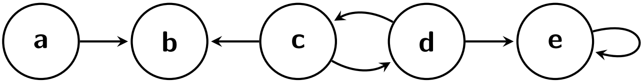
A (three–valued) labeling is a total function $Lab : A \rightarrow \{in, out, undec\}$. We can write it as tuple (I, O, U) where $I/O/U$ stand for sets of arguments mapped respectively to $in, out, undec$.
An *in*–labeled argument is **legally** *in* iff all its attackers are labeled *out*.
An *out*–labeled argument is **legally** *out* iff at least one its attacker is labeled *in*.

Note: sometimes one can also use $\{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$ instead of $\{in, out, undec\}$.

Labeling–based definition

Labeling *Lab* is admissible in *AF* iff each *in*–labeled argument is legally *in* and each *out*–labeled argument is legally *out*.

Example



Admissible extensions: $\{a, c\}$,$\{a, d\}$, $\{a\}$, $\{c\}$, $\{d\}$ and $\emptyset$.
Admissible labellings: $(\{a, c\}, \{d\}, \{b, e\})$, $(\{a, c\}, \{b, d\}, \{e\})$, $(\{a, d\}, \{c\}, \{b, e\})$, $(\{a, d\}, \{b, c\}, \{e\})$, $(\{a, d\}, \{c, e\}, \{b\})$, $(\{a, d\}, \{b, c, e\}, \emptyset)$, $(\{a\}, \emptyset, \{b, c, d, e\}$, $(\{a\}, \{b\}, \{c, d, e\}$, $(\{c\}, \{d\}, \{a, b, e\}$, $(\{c\}, \{b, d\}, \{a, e\}$, $(\{d\}, \{c\}, \{a, b, e\}$, $(\{d\}, \{c, e\}, \{a, b\}$, $(\emptyset, \emptyset, \{a, b, c, d, e\})$

Comparison

Range

By S^+ we understand the **set of arguments attacked** by S. We will also refer to it as the **discarded set**. The set $S \cup S^+$ is called the **range** of S.

The two approaches are **equivalent**:

Extension–based to labeling–based

If S is an admissible extension, then $(S, S^+, A \setminus (S \cup S^+)$ is an admissible labeling.

Labeling–based to extension–based

If Lab is an admissible labeling, then $in(Lab)$ is an admissible extension.

Abstract Dialectical Framework I

Definition

An **abstract dialectical framework** (ADF) is a tuple (S, L, C), where:

- S is a set of abstract **arguments** (nodes, statements),
- $L \subseteq S \times S$ is a set of **links** (edges) and
- $C = \{C_s\}_{s \in S}$ is a set of **acceptance conditions**, one condition per each argument.

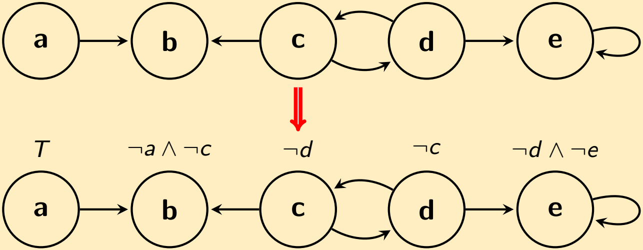
Important: links now do not represent relations anymore; the precise nature of the interaction between arguments is specified by the acceptance conditions.

Acceptance conditions

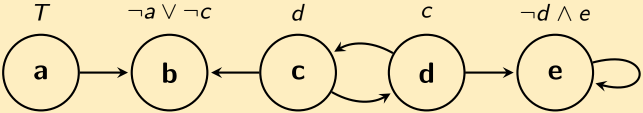
- They represent the relation of an argument to its parents
- Can be represented as functions $C_s : 2^{par(s)} \rightarrow \{in, out\}$
- More commonly defined as propositional formulas

Abstract Dialectical Framework II

Example



Example



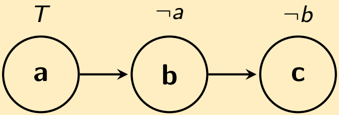
The idea behind admissible semantics I

What is admissible semantics all about?

- An admissible set of arguments can **stand on its own** [1], i.e. it can respond with attacks to incoming attacks
- A dialog view: whatever our opponent utters against us, we can provide some sort of a counterargument to it

Question is: how to make sure that we properly discard the "undesired" arguments?

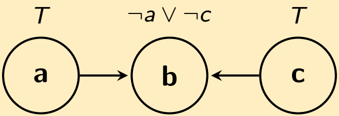
Examples



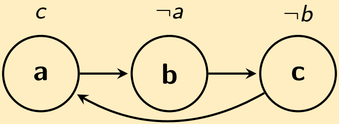
Is $\{c\}$ admissible? How about $\{b\}$?

The idea behind admissible semantics II

Example



Is $\{b\}$ admissible?



Is $\{b\}$ admissible? Is $\{a, c\}$ admissible?

[1] P. Baroni and G. Mssimiliano, Semantics of Abstract Argument Systems, Argumentation in AI 25-44, 2009

Basic notions

Let us assume an ADF $D = (S, L, C)$. By $par(a)$ we understood the set of parents of an argument a.

Conflict–freeness

A set $A \subseteq S$ is **conflict–free** in D iff for every argument $a \in A$ its acceptance condition is met, i.e. $C_a(A \cap par(a)) = in$ where $par(a)$.

Idea: the outcomes of extension–based and labeling–based semantics can be uniformly presented as two or three–valued interpretations.

Interpretation

Interpretations (labellings) map arguments to truth values. We say v is two–valued if $v : S \to \{\mathbf{t}, \mathbf{f}\}$ and three valued if $v : S \to \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$. We say that an interpretation is **partial** if it is defined on a set $A \subseteq S$.

By $v^{\mathbf{t}}$ we will denote the set of arguments mapped to **t** by v, similarly for $v^{\mathbf{f}}$ and $v^{\mathbf{u}}$.

Extension–based admissibility in ADFs I

Extension based approach: the idea

Assume we have a (conflict–free) set of arguments A. How to check if it is admissible?

- Identify what arguments "attack" us:
  We need to "protect" a set of arguments F if for any $a \in A$, it **changes its acceptance condition to** *out*: $C_a(par(a) \cap (A \cup F)) = out$.
- How to make sure that F cannot be uttered?
  We need to "attack" any of its elements back: we achieve that by making sure that **part of** F **is contained in the discarded set** A^+.

Problem: how to properly compute A^+ in an abstract setting?

Extension–based admissibility in ADFs II

Completion

By a **completion** of an interpretation v to a set Z where $A \subseteq Z$ we understand an interpretation v' defined on Z in a way that $\forall a \in Z \setminus A \; v(a) = v'(a)$.
We say that v' is a **t/f/u–completion** if it maps all elements from $Z \setminus A$ respectively to **t**/**f**/**u**.

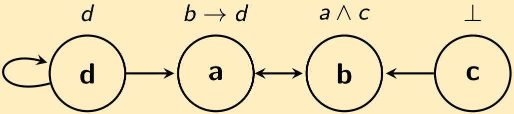
Idea: partial interpretation can be enough to get the "final" value of a formula.

Decisive interpretation

Let v_Z be a two or three–valued interpretation defined on a set $Z \subseteq S$, We say that v_Z is **decisive** for s iff for any two (respectively two or three–valued) completions $v_{par(s)}$ and $v'_{par(s)}$ of v to $Z \cup par(s)$, it holds that $v_{par(s)}(C_s) = v'_{par(s)}(C_s)$.
We say that s is **decisively out/in/undecided** wrt v_Z if v_Z is decisive and all of its completions map s to respectively *out*, *in*, *undec*.

Extension–based admissibility in ADFs III

Example



Example of a decisive interpretation for a: $v = \{b : \mathbf{f}\}$
Example of a decisive interpretation for b: $v = \{c : \mathbf{f}\}$

Extension–based admissibility in ADFs IV

Range and discarded set

Let $A \subseteq S$ be a conflict–free extension of D. Let v be a partial two–valued interpretation built as follows:

1. Let $M = A$. For every $a \in M$ set $v(a) = \mathbf{t}$.
2. For every argument $b \in S \setminus M$ that is decisively out in v, set $v(b) = \mathbf{f}$ and add b to M.
3. Repeat the previous step until there are no new elements added to M.

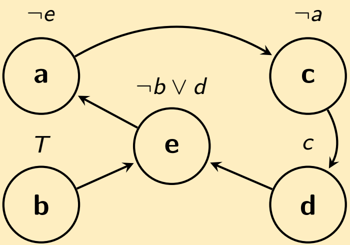
By A^+ we understand the set of arguments mapped to $\mathbf{f}$ by v. The **range** of A, denoted A^R is defined as $A \cup A^+$. We refer to v as **range interpretation** of A.

Extension–based admissibility in ADFs V

Decisive admissibility

Let $A \subseteq S$ be a conflict–free extension of D and A^+ its discarded set. A is **admissible** in D iff for any $F \subseteq S \setminus A$ ($F \neq \emptyset$), if there exists an $a \in A$ s.t. $C_a(par(a) \cap (F \cup A)) = out$ then $F \cap A^+ \neq \emptyset$.

Example



Is $A = \{a\}$ admissible?
The discarded set A^+ is $\{c, d\}$. All the "attacking sets" revolve around e, which is not contained in A^+. Hence, NO.
Is $A = \{a, b\}$ admissible?
The discarded set A^+ is $\{c, d, e\}$. This time e is contained. The set is admissible.

Labeling–based admissibility in ADFs I

An approach introduced by Brewka, Ellmauthaler, Strass, Wallner and Woltran in "Abstract Dialectical Framework Revisited", IJCAI 2013.

Labeling–based approach: the idea

Assume a three–valued interpretation v. How to check if it is admissible?

- The value of the acceptance condition has to agree with the argument assignment: conflict–freeness
- Imagine the arguments mapped to **u** as wildcards: if changing the assignment **t** or **f** "breaks" our interpretation, it means there are counterarguments we cannot protect ourselves against

Core of the solution: check if all possible interpretations without "wildcards" **agree with each other**

Labeling–based admissibility in ADFs II

Comparing interpretations

We can compare interpretations according to how much **information** they carry using the ordering $\leq_i$. According to it **u** carries the least information, i.e. $\mathbf{u} \leq_i \mathbf{f}$ and $\mathbf{u} \leq_i \mathbf{t}$.
It can be generalized for interpretations: given two interpretations v and v' defined on S we say that $v \leq_i v'$ iff $\forall s \in S$ $v(s) \leq_i v'(s)$.

Two–valued extensions of an interpretation

Given a three valued interpretation v, by $[v]_2$ we understand the set of all two–valued interpretations w s.t. $v \leq_i w$.

In other words: set of all interpretations with each **u** substituted by **t** or **f**.

Labeling–based admissibility in ADFs III

Consensus

The operator $\sqcap$ represents the **consensus** among values: $\mathbf{t} \sqcap \mathbf{t} = \mathbf{t}$, $\mathbf{f} \sqcap \mathbf{f} = \mathbf{f}$ and $\mathbf{u}$ in all other cases.

Characteristic operator

Let v a three–valued interpretation defined over S, s an argument in S and $\Gamma_D : (S \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}) \rightarrow (S \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{u}\})$ a function from three–valued interpretations to three–valued interpretations. Then $\Gamma_D(v) = v'$ with

$$v'(s) = \bigsqcap_{w \in [v]_2} C_s(par(s) \cap w^{\mathbf{t}})$$

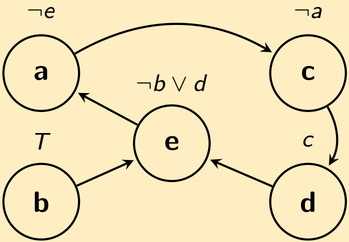
Labeling–based admissibility in ADFs IV

Three–valued admissibility

A three-valued interpretation v is **admissible** in D iff $v \leq_i \Gamma_D(v)$.

Example

Is $v = (\{a, b\}, \{c, d\}, \{e\})$ an admissible interpretation?

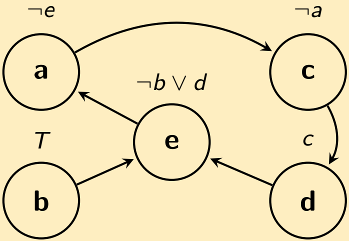


a	b	c	d	e
$[v]_2$				
t	**t**	**f**	**f**	**t**
t	**t**	**f**	**f**	**f**
New values of C_s				
f	**t**	**f**	**f**	**f**
t	**t**	**f**	**f**	**f**
Consensus				
u	**t**	**f**	**f**	**f**

Labeling–based admissibility in ADFs V

Example

Is $v = (\{a, b\}, \{c, d, e\}, \emptyset)$ **an admissible interpretation?**



a	b	c	d	e
$[v]_2$				
t	**t**	**f**	**f**	**f**
New values of C_s				
t	**t**	**f**	**f**	**f**
Consensus				
t	**t**	**f**	**f**	**f**

Comparison

Although the motivations behind the two approaches is similar, they are in general **not equivalent**.

Decisive to three–valued

If A is a decisive admissible extension, then $(A, A^+, S \setminus (A \cup A^+)$ is a three–valued admissible interpretation.

Three–valued to decisive

Let $D = (S, L, C)$ be a bipolar ADF without support cycles and v a lattice admissible three-valued interpretation in D. Then $A = v^{\mathbf{t}}$ is a decisive admissible extension of D.

Discussion

- Is there one good approach for admissibility?
- Is the nonequivalence between formulations in ADFs a sign of error?
- Is there an agreement on the semantics between other generalizations of Dung framework?


this is how I finish a presentation:
happymonsters.tumblr.com
soo...
uhh
...yeah.