Weighted Abstract Dialectical Frameworks

Gerhard Brewka

Computer Science Institute
University of Leipzig
brewka@informatik.uni-leipzig.de

joint work with H. Strass, J. Wallner, S. Woltran

1. Motivation

- Dung frameworks widely used in abstract argumentation
- Nice and simple tool, yet restricted to expressing attack
- Various generalizations including other relations (e.g. support)
- ADFs provide a systematic generalization
- but still lack facilities to express argument strength
- This is what we want to add today

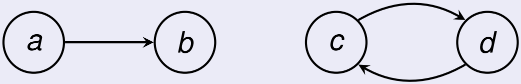
Outline

1. Motivation (done)
2. AFs: A Reconstruction
3. From AFs to ADFs
4. From ADFs to Weighted ADFs
5. Alternative Valuation Structures
6. Conclusion

2. Dung Frameworks

Abstract Argumentation Frameworks (AFs)

- syntactically: directed graphs



- conceptually: nodes are arguments, edges denote attacks between arguments
- semantically: *extensions* are sets of “acceptable” arguments
- immensely popular in the argumentation community

AF Semantics

Let $F = (A, R)$ be an argumentation framework, $S \subseteq A$.

- S is *conflict-free* iff no element of S attacks an element in S.
- $a \in A$ is *defended* by S iff all attackers of a are attacked by an element of S.
- a conflict-free set S is
  - *admissible* iff it defends all arguments it contains,
  - *preferred* iff it is $\subseteq$-maximal admissible,
  - *complete* iff it contains exactly the arguments it defends,
  - *grounded* iff it is $\subseteq$-minimal complete,
  - *stable* iff it attacks all arguments not in S.

Operator-based Reconstruction

- S splits arguments into subsets: in S, attacked by S, undefined.
- Calls for analysis in terms of partial interpretations.
- Based on operator Γ_D over partial interpretations (here represented as consistent sets of literals).
- Takes interpretation v and produces a new (revised) one v'.
- $v' = \Gamma_D(v)$ makes a node s
  - **t** iff s unattacked in all 2-valued completions of v,
  - **f** iff s attacked in all 2-valued completions of v,
  - undefined otherwise.
- Operator thus checks what can be justified based on v.

Operator-based Reconstruction

- S splits arguments into subsets: in S, attacked by S, undefined.
- Calls for analysis in terms of partial interpretations.
- Based on operator Γ_D over partial interpretations (here represented as consistent sets of literals).
- Takes interpretation v and produces a new (revised) one v'.
- $v' = \Gamma_D(v)$ makes a node s
  - **t** iff s unattacked in all 2-valued completions of v,
  - **f** iff s attacked in all 2-valued completions of v,
  - undefined otherwise.
- Operator thus checks what can be justified based on v.

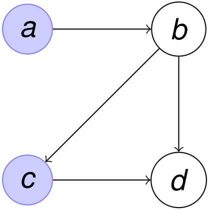
Operator Based Semantics

An interpretation v of an AF D is

- a *model* of D iff v is two-valued and $\Gamma_D(v) = v$.
  Intuition: argument is **t** iff no attacker is **t**.
- *grounded* for D iff it is the least fixpoint of Γ_D.
  Intuition: collects information beyond doubt.
- *admissible* for D iff $v \subseteq \Gamma_D(v)$
  Intuition: does not contain unjustifiable information
- *preferred* for D iff it is $\subseteq$-maximal admissible for D
  Intuition: want maximal information content.
- *complete* for D iff $v = \Gamma_D(v)$.
  Intuition: contains exactly the justifiable information.

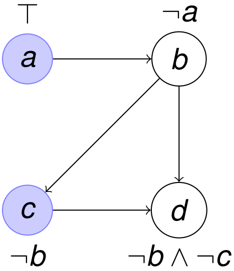
Result: Dung extensions $\Longleftrightarrow$ arguments **t** in respective interpretations.

3. ADFs: Basic Idea



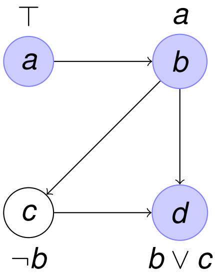
An Argumentation Framework

Basic Idea



An Argumentation Framework
with explicit acceptance conditions

Basic Idea



A Dialectical Framework
with flexible acceptance conditions

ADF Definition

Syntax

Definition: Abstract Dialectical Framework

An abstract dialectical framework (ADF) is a triple $D = (S, L, C)$,

- S . . . set of statements, arguments; anything one might accept
- $L \subseteq S \times S$ . . . links
- $C = \{\varphi_s\}_{s \in S}$ . . . acceptance conditions

- links denote a dependency
- acceptance condition: defines truth value for s based on truth values of its parents
- specified as propositional formula φ_s

ADFs: Semantics

Analysis in terms of partial interpretations; handle on what is unknown.

Truth values, interpretations

- truth values: true **t**, false **f**; **u** stands for undefined
- partial interpretation: $v : S \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$
- interpretations can be represented as consistent sets of literals

Information ordering

- $\mathbf{u} <_i \mathbf{t}$ and $\mathbf{u} <_i \mathbf{f}$ (as usual $x \leq_i y$ iff $x <_i y$ or $x = y$)
- *consensus* $\sqcap$ is greatest lower bound w.r.t. $\leq_i$:
  $\mathbf{t} \sqcap \mathbf{t} = \mathbf{t}$ and $\mathbf{f} \sqcap \mathbf{f} = \mathbf{f}$, otherwise $x \sqcap y = \mathbf{u}$
- information ordering generalised to interpretations:
  $v_1 \leq_i v_2$ iff $v_1(s) \leq_i v_2(s)$ for all $s \in S$

The Characteristic Operator

- Takes interpretation v and produces a new (revised) one v'.
- v' makes a node s
  - **t** iff acceptance condition true under any 2-valued completion of v,
  - **f** iff acceptance condition false under any 2-valued completion of v,
  - **u** otherwise.
- Operator thus checks what can be justified based on v.
  - Can information in v be justified?
  - Can further information be justified?

Characteristic Operator Γ_D

- for interpretation v, we define $[v]_2 = \{v \leq_i w \mid w \text{ two-valued}\}$
- for interpretation $v : S \to \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$, Γ_D yields a new interpretation (the consensus over $[v]_2$)

$$\Gamma_D(v) : S \to \{\mathbf{t}, \mathbf{f}, \mathbf{u}\} \qquad s \mapsto \bigsqcap \{w(\varphi_s) \mid w \in [v]_2\}$$

Semantics via Fixed Points

A partial interpretation v of ADF D is

- a *model* of D iff v is two-valued and $\Gamma_D(v) = v$.
  Intuition: statement is **t** iff its acceptance condition says so.
- *grounded* for D iff it is the least fixpoint of Γ_D.
  Intuition: collects information beyond doubt.
- *admissible* for D iff $v \leq_i \Gamma_D(v)$
  Intuition: does not contain unjustifiable information
- *preferred* for D iff it is $\leq_i$-maximal admissible for D
  Intuition: want maximal information content.
- *complete* for D iff $v = \Gamma_D(v)$.
  Intuition: contains exactly the justifiable information.

Stable Models for ADFs

Based on ideas from Logic Programming:

- no self-justifying cycles,
- achieved by reduct-based check.

To check whether a two-valued model v of D is *stable* do the following:

- eliminate in D all nodes with value **f** and corresponding links,
- replace eliminated nodes in acceptance conditions by **f**,
- check whether nodes **t** in v coincide with grounded model of reduced ADF.

Stable Models for ADFs

Based on ideas from Logic Programming:

- no self-justifying cycles,
- achieved by reduct-based check.

To check whether a two-valued model v of D is *stable* do the following:

- eliminate in D all nodes with value **f** and corresponding links,
- replace eliminated nodes in acceptance conditions by **f**,
- check whether nodes **t** in v coincide with grounded model of reduced ADF.

Results

- ADFs properly generalize AFs.
- All major semantics available.
- Many results carry over, eg. the following inclusions hold:

$$sta(D) \subseteq val_2(D) \subseteq pref(D) \subseteq com(D) \subseteq adm(D).$$

- for ADFs corresponding to AFs models and stable models coincide (as AFs cannot express support).
- Complexity increases by one level in PH as compared to AFs.
- Stays the same for interesting subclass of bipolar ADFs.

Results

- ADFs properly generalize AFs.
- All major semantics available.
- Many results carry over, eg. the following inclusions hold:

$$sta(D) \subseteq val_2(D) \subseteq pref(D) \subseteq com(D) \subseteq adm(D).$$

- for ADFs corresponding to AFs models and stable models coincide (as AFs cannot express support).
- Complexity increases by one level in PH as compared to AFs.
- Stays the same for interesting subclass of bipolar ADFs.

Results

- ADFs properly generalize AFs.
- All major semantics available.
- Many results carry over, eg. the following inclusions hold:

$$sta(D) \subseteq val_2(D) \subseteq pref(D) \subseteq com(D) \subseteq adm(D).$$

- for ADFs corresponding to AFs models and stable models coincide (as AFs cannot express support).
- Complexity increases by one level in PH as compared to AFs.
- Stays the same for interesting subclass of bipolar ADFs.

Results

- ADFs properly generalize AFs.
- All major semantics available.
- Many results carry over, eg. the following inclusions hold:

$$sta(D) \subseteq val_2(D) \subseteq pref(D) \subseteq com(D) \subseteq adm(D).$$

- for ADFs corresponding to AFs models and stable models coincide (as AFs cannot express support).
- Complexity increases by one level in PH as compared to AFs.
- Stays the same for interesting subclass of bipolar ADFs.

4. Weighted ADFs Over [0, 1]

Major change:
Acceptance degrees of arguments taken from unit interval [0, 1].

Syntax:

Definition: Weighted ADF

A weighted ADF (WADF) is a triple $D = (S, L, C)$,

- S ... set of statements, arguments; anything one might accept
- $L \subseteq S \times S$ ... links
- $C = \{\varphi_s\}_{s \in S}$ ... acceptance conditions

- basically unchanged, but:
- elements of [0, 1] allowed as constants in formulas φ_s
- allows us to fix acceptance degrees, and to express upper and lower bounds, e.g.($\phi \wedge 0.7$), respectively ($\phi \vee 0.7$)

Weighted ADFs: Semantics

Interpretions of interest: partial functions from S to $[0, 1]$,
that is, functions $v : S \rightarrow [0, 1] \cup \{\mathbf{u}\}$

Completions: replace value **u** by values in $[0, 1]$ in all possible ways, rest untouched

2 remaining questions:

- Q: How to evaluate acceptance conditions? Possible answer:
- Acceptance degrees evaluate to themselves;
  $\wedge$ to *min*, $\vee$ to max, $\neg y$ to $1 - y$
- Q: What's the information ordering? Possible answer:
- $\mathbf{u} <_i x$ for all $x \in [0, 1]$, all other values incomparable

The Characteristic Operator

- Takes partial interpretation v and produces a new one v'.
- $v'(s)$ is the glb wrt. $<_i$ (consensus) of the set

$$\{w(\varphi_s) \mid w \text{ completion of } v\}$$

- the rest falls into place

Characteristic Operator Γ_D

Let $v : S \to [0, 1] \cup \{\mathbf{u}\}$ be a partial interpretation

- define $[v]_c = \{w \mid w \text{ completion of } v\}$
- Γ_D yields a new interpretation (the consensus over $[v]_c$)

$$\Gamma_D(v) : S \to [0, 1] \cup \{\mathbf{u}\} \qquad s \mapsto \bigsqcap \{w(\varphi_s) \mid w \in [v]_c\}$$

Semantics via Fixed Points

A partial interpretation of WADF D is

- a *model* of D iff $v(s) \neq \mathbf{u}$ for all $s \in S$ and $\Gamma_D(v) = v$.
  Intuition: value of s is the one required by acceptance condition.
- *grounded* for D iff it is the least fixpoint of Γ_D.
  Intuition: collects information beyond doubt.
- *admissible* for D iff $v \leq_i \Gamma_D(v)$
  Intuition: does not contain unjustifiable information
- *preferred* for D iff it is $\leq_i$-maximal admissible for D
  Intuition: want maximal information content.
- *complete* for D iff $v = \Gamma_D(v)$.
  Intuition: contains exactly the justifiable information.

A Caveat

- Need to show monotonicity of operator Γ_D (otherwise grounded not well-defined)
- Monotonic wrt. extension of $<_i$ to interpretations (componentwise)
- Follows from the fact that $v_1 \leq_i v_2$ implies $[v_2]_c \subseteq [v_1]_c$

Relation to ADFs

- D WADF, no constants other than 0, 1 in acceptance conditions; v partial interpretation assigning truth values in $\{0, 1, \mathbf{u}\}$ only:
  - if $\Gamma_D^{WDF}(v)(s) = 1/0$, then so is if $\Gamma_D^{ADF}(v)(s)$
  - if $\Gamma_D^{ADF}(v)(s) = \mathbf{u}$, then so is if $\Gamma_D^{WDF}(v)(s)$
  - but $\Gamma_D^{ADF}(v)(s)$ can be 1/0, yet $\Gamma_D^{WDF}(v)(s) = \mathbf{u}$
- Example: nodes a, b with acceptance conditions $a : a$; $b : a \vee \neg a$
  - grounded (weighted): $a \to \mathbf{u}$, $b \to \mathbf{u}$
  - in standard approach: $a \to \mathbf{u}$, $b \to 1$
- Possible solution: define “strange" connectives:
  - conjunction: $x \wedge y = 1$ iff $x > 0.5$ and $y > 0.5$, 0 otherwise
  - disjunction: $x \vee y = 1$ iff $x > 0.5$ or $y > 0.5$, 0 otherwise

Alternative Options

- Alternative evaluation of connectives, e.g. Łukasiewicz:
  - strong conjunction: $x \tilde{\wedge} y = max\{0, x + y - 1\}$
  - strong disjunction: $x \tilde{\vee} y = min\{1, x + y\}$
- More refined information ordering, e.g. **u** $<_i$ 0.5 and additionally

  $0 >_i ... >_i 0.1 ... >_i 0.4 >_i ... >_i 0.5 <_i ... <_i 0.6 <_i ... <_i 0.9 ... <_i 1$

- Handle on which interpretations are admissible/preferred

5. Alternative Valuation Structures

- So far only considered values in $[0, 1]$
- Many more options, e.g.:
  - $W_m = \{\frac{k}{m-1} \mid 0 \leq k \leq m - 1\}$, e.g. $W_5 = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$
  - Belnap's 4-valued system with $\{\emptyset, \{\bot\}, \{\top\}, \{\bot, \top\}\}$
  - intervals from within $[0, 1]$
- Interpretations assign value from chosen structure to nodes in S
- Need to define connectives and information ordering
- Literature on multi-valued logics offers wide range of options
  - valuation structures
  - evaluations of propositional formulas (Gödel, Łukasiewicz, etc.)
- Information ordering semi-lattice with least element **u**

Example 1: W_3

- truth degrees: $\{0, 0.5, 1\}$
- formula evaluation as before: 0, 0.5, 1 evaluate to themselves; $\wedge$ to *min*, $\vee$ to max, $\neg y$ to $1 - y$
- information ordering: $\mathbf{u} <_i x$ for all $x \in W_3$, rest incomparable
- and we're done!
- note that $0.5 \neq \mathbf{u}$:
  acceptance degree is 0.5 vs. acceptance degree is unknown

Example 2: Belnap

- truth degrees: $\{\emptyset, \{\bot\}, \{\top\}, \{\bot, \top\}\}$
- formula evaluation:
  conjunction/disjunction are inf/sup under truth ordering;
  negation swaps $\{\bot\}$ and $\{\top\}$, leaves other values unchanged
- truth ordering:

$$\{\bot\} <_t \emptyset <_t \{\top\} \qquad \{\bot\} <_t \{\bot, \top\} <_t \{\top\}$$

- information ordering:

$$\mathbf{u} <_i \emptyset <_i \{\top\} <_i \{\bot, \top\} \qquad \mathbf{u} <_i \emptyset <_i \{\bot\} <_i \{\bot, \top\}$$

- and we're done;
  note that $\mathbf{u} \neq \emptyset$; treating them as identical yields different system

Example 3: Intervals

- truth degrees: $INT = \{[a, b] \mid 0 \leq a \leq b \leq 1\}$
- formula evaluation:

$$[a, b] \wedge [c, d] = [min(a, c), min(b, d)]$$
$$[a, b] \vee [c, d] = [max(a, c), max(b, d)]$$
$$\neg[a, b] = [1 - b, 1 - a]$$

- information ordering: $\mathbf{u} <_i v$ for all $v \in INT$, in addition

$$[a, b] <_i [c, d] \text{ iff } [c, d] \subsetneq [a, b]$$

- characteristic operator now fully defined

6. Conclusions

- Presented weighted ADFs, a generalization of ADFs allowing to assign acceptance degrees to arguments
- Introduced approach using values from unit interval
- Showed what needs to be done to get ADF techniques to work
  - Define formula evaluation
  - Define information ordering
- Illustrated how alternative valuation structures can be integrated
- Paves the way to bridge multi-valued logics with argumentation: e.g. $ASPIC^{+}_{mv}$, a multi-valued variant of *ASPIC*?
- Wide range of options: which are the interesting ones?

6. Conclusions

- Presented weighted ADFs, a generalization of ADFs allowing to assign acceptance degrees to arguments
- Introduced approach using values from unit interval
- Showed what needs to be done to get ADF techniques to work
  - Define formula evaluation
  - Define information ordering
- Illustrated how alternative valuation structures can be integrated
- Paves the way to bridge multi-valued logics with argumentation: e.g. $ASPIC^{+}_{mv}$, a multi-valued variant of *ASPIC*?
- Wide range of options: which are the interesting ones?

6. Conclusions

- Presented weighted ADFs, a generalization of ADFs allowing to assign acceptance degrees to arguments
- Introduced approach using values from unit interval
- Showed what needs to be done to get ADF techniques to work
  - Define formula evaluation
  - Define information ordering
- Illustrated how alternative valuation structures can be integrated
- Paves the way to bridge multi-valued logics with argumentation: e.g. $ASPIC^{+}_{mv}$, a multi-valued variant of *ASPIC*?
- Wide range of options: which are the interesting ones?

6. Conclusions

- Presented weighted ADFs, a generalization of ADFs allowing to assign acceptance degrees to arguments
- Introduced approach using values from unit interval
- Showed what needs to be done to get ADF techniques to work
  - Define formula evaluation
  - Define information ordering
- Illustrated how alternative valuation structures can be integrated
- Paves the way to bridge multi-valued logics with argumentation: e.g. $ASPIC^{+}_{mv}$, a multi-valued variant of $ASPIC$?
- Wide range of options: which are the interesting ones?

6. Conclusions

- Presented weighted ADFs, a generalization of ADFs allowing to assign acceptance degrees to arguments
- Introduced approach using values from unit interval
- Showed what needs to be done to get ADF techniques to work
  - Define formula evaluation
  - Define information ordering
- Illustrated how alternative valuation structures can be integrated
- Paves the way to bridge multi-valued logics with argumentation: e.g. $ASPIC^{+}_{mv}$, a multi-valued variant of *ASPIC*?
- Wide range of options: which are the interesting ones?

6. Conclusions

- Presented weighted ADFs, a generalization of ADFs allowing to assign acceptance degrees to arguments
- Introduced approach using values from unit interval
- Showed what needs to be done to get ADF techniques to work
  - Define formula evaluation
  - Define information ordering
- Illustrated how alternative valuation structures can be integrated
- Paves the way to bridge multi-valued logics with argumentation: e.g. $ASPIC^{+}_{mv}$, a multi-valued variant of *ASPIC*?
- Wide range of options: which are the interesting ones?

6. Conclusions

- Presented weighted ADFs, a generalization of ADFs allowing to assign acceptance degrees to arguments
- Introduced approach using values from unit interval
- Showed what needs to be done to get ADF techniques to work
  - Define formula evaluation
  - Define information ordering
- Illustrated how alternative valuation structures can be integrated
- Paves the way to bridge multi-valued logics with argumentation: e.g. $ASPIC^{+}_{mv}$, a multi-valued variant of *ASPIC*?
- Wide range of options: which are the interesting ones?