Bridges between Abstract Argumentation and Belief Revision

Sylvie Coste-Marquis Sébastien Konieczny
Jean-Guy Mailly Pierre Marquis

Centre de Recherche en Informatique de Lens
Université d'Artois – CNRS UMR 8188

2nd Madeira Workshop on Belief Revision and Argumentation
February 9th – February 13th

AMANDE

Outline

Outline

Introduction
- Abstract Argumentation
- Belief Revision

Overview of our Contributions
- Adapting AGM to Abstract Argumentation
- Using AGM to Revise Abstract AF

Translation-based Revision of Argumentation Frameworks
- Encoding AF and their Semantics
- Distance-based Operators and Minimal Change

Conclusion and Future Work

Abstract Argumentation [Dung 1995]

- An abstract argumentation framework is a pair $\langle \mathcal{A}, \mathcal{R} \rangle$ with $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$:

- An extension is a set of arguments that can be accepted together
  - Different semantics to define the extensions: complete, stable, preferred, grounded, etc.
- The aim is to know whether an argument is accepted or not w.r.t. the chosen semantics σ
  - An argument $a \in \mathcal{A}$ is (skeptically) accepted iff it belongs to every extension of the AF w.r.t. the considered semantics σ:

$$F \mid\!\sim_{\sigma} a \Leftrightarrow a \in \bigcap Ext_{\sigma}(F)$$

AGM Framework for Belief Revision

- AGM Framework [Alchourrón, Gärdenfors and Makinson 1985]
- Adaptation for propositional logic [Katsuno and Mendelzon 1991]
- Incorporate a new piece of information α in the agent's beliefs φ wrt some notion of plausibility p:

$$Mods(\varphi \circ \alpha) = \min(Mods(\alpha), \leq_p)$$

AMANDE

UNIVERSITÉ D'ARTOIS

AF Revision

- Aim: Incorporation of a new piece of information about the attack relation and/or the acceptance statuses of arguments
- Two kind of minimal change:

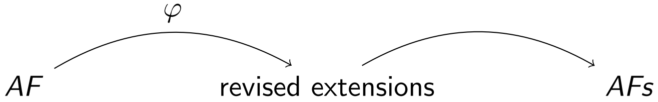
$$\text{Attack} \neq \text{Acceptance}$$

Outline

AMANDE

Adapting AGM to Abstract Argumentation

- A Two-step Process



Summary of this Contribution

- New piece of information: formula about acceptance statuses
ex: $\varphi = (a_1 \vee a_2) \wedge \neg a_3$
- First minimality criterion: minimal change of arguments statuses
- Other (less important) minimality criterion: minimal change of the attack relation, minimality of the output's size
- More details: Coste-Marquis, Konieczny, Mailly, Marquis, *On the Revision of Argumentation Systems: Minimal Change of Arguments Statuses*, KR 2014

Using AGM to Revise Abstract AF

- σ: a semantics to define acceptable arguments
- F: an argumentation framework
- φ: a propositional formula indicating how to revise F

F, φ

Using AGM to Revise Abstract AF

- σ: a semantics to define acceptable arguments
- F: an argumentation framework
- φ: a propositional formula indicating how to revise F

F, φ

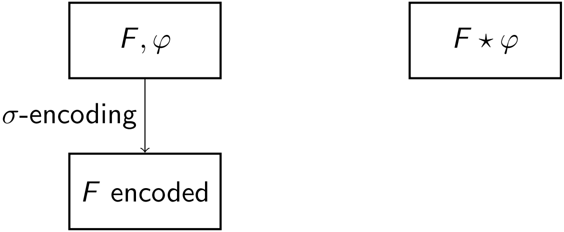
$F \star \varphi$

AMANDE

UNIVERSITÉ D'ARTOIS

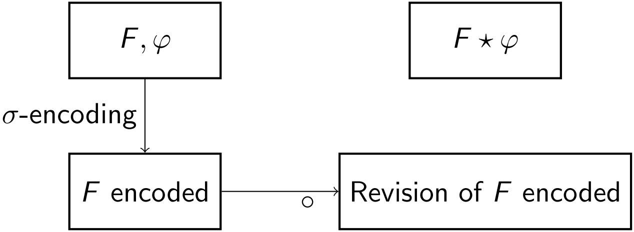
Using AGM to Revise Abstract AF

- σ: a semantics to define acceptable arguments
- F: an argumentation framework
- φ: a propositional formula indicating how to revise F



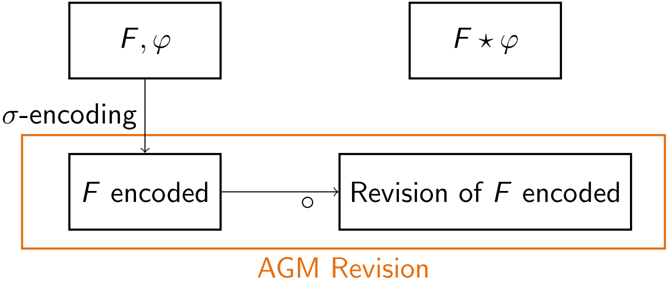
Using AGM to Revise Abstract AF

- σ: a semantics to define acceptable arguments
- F: an argumentation framework
- φ: a propositional formula indicating how to revise F



Using AGM to Revise Abstract AF

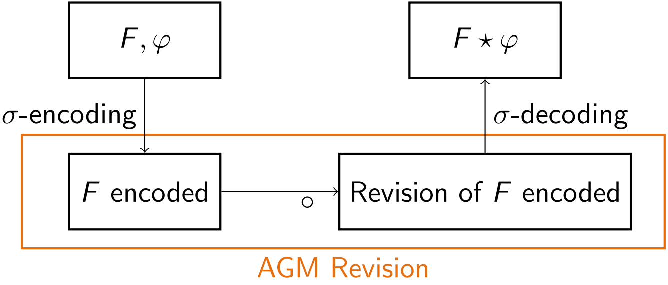
- σ: a semantics to define acceptable arguments
- F: an argumentation framework
- φ: a propositional formula indicating how to revise F



AMANDE

Using AGM to Revise Abstract AF

- σ: a semantics to define acceptable arguments
- F: an argumentation framework
- φ: a propositional formula indicating how to revise F

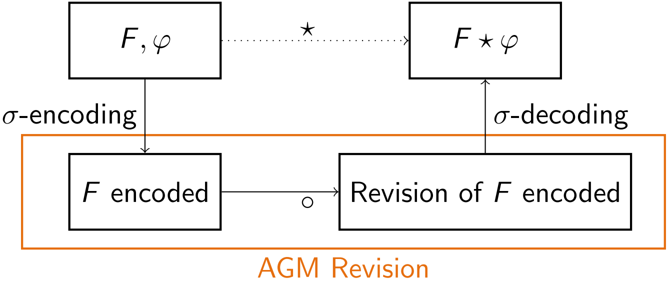


AMANDE

UNIVERSITÉ D'ARTOIS

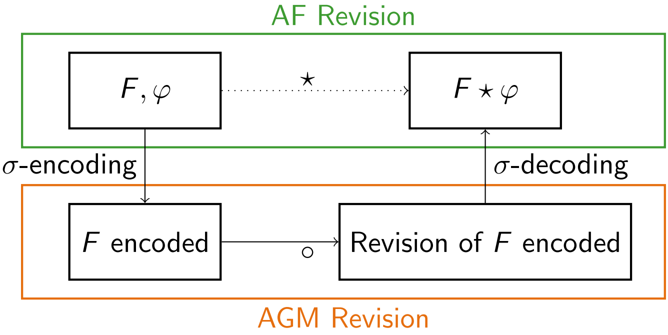
Using AGM to Revise Abstract AF

- σ: a semantics to define acceptable arguments
- F: an argumentation framework
- φ: a propositional formula indicating how to revise F



Using AGM to Revise Abstract AF

- σ: a semantics to define acceptable arguments
- F: an argumentation framework
- φ: a propositional formula indicating how to revise F



Outline

Propositional Language

- $\forall x \in A$, $acc(x) =$ " x is skeptically accepted by F "
- $\forall x, y \in A$, $att(x, y) =$ " x attacks y in F "
- $Prop_A = \{acc(x) | x \in A\} \cup \{att(x, y) | x, y \in A\}$
- $\mathcal{L}_A$ is the propositional language built on the set of variables $Prop_A$ and the connectives $\neg, \vee, \wedge$

Encoding an AF

σ-formula of F

Given an AF $F = \langle A, R \rangle$ and a semantics σ, the σ-formula of F is

$$f_\sigma(F) = \bigwedge_{(x,y)\in R} att(x,y) \wedge \bigwedge_{(x,y)\notin R} \neg att(x,y)$$

Encoding an AF

σ-formula of F

Given an AF $F = \langle A, R \rangle$ and a semantics σ, the σ-formula of F is

$$f_\sigma(F) = \bigwedge_{(x,y)\in R} att(x,y) \wedge \bigwedge_{(x,y)\notin R} \neg att(x,y) \wedge th_\sigma(A)$$

where the σ-theory of A $th_\sigma(A)$ is a formula which encodes the semantics σ.

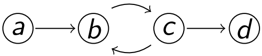
Encoding the Stable Semantics (1)

Stable extensions of an AF $F = \langle A, R \rangle$ [Besnard and Doutre 2004]

$$\bigwedge_{a \in A} \left(a \Leftrightarrow \bigwedge_{b:(b,a) \in R} \neg b\right)$$

Example

$F_1 =$



$$a \wedge [b \Leftrightarrow (\neg a \wedge \neg c)]$$
$$\wedge [c \Leftrightarrow \neg b] \wedge [d \Leftrightarrow \neg c]$$

One single model / stable extension: $\{a, c\}$

AMANDE

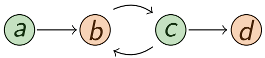
UNIVERSITÉ D'ARTOIS

Encoding the Stable Semantics (1)

Stable extensions of an AF $F = \langle A, R \rangle$ [Besnard and Doutre 2004]

$$\bigwedge_{a \in A} (a \Leftrightarrow \bigwedge_{b:(b,a) \in R} \neg b)$$

Example

$F_1 =$ 

$$a \wedge [b \Leftrightarrow (\neg a \wedge \neg c)]$$
$$\wedge [c \Leftrightarrow \neg b] \wedge [d \Leftrightarrow \neg c]$$

One single model / stable extension: $\{a, c\}$

CNRS cril AMANDE UNIVERSITÉ D'ARTOIS

Encoding the Stable Semantics (2)

From

$$\bigwedge_{a \in A} (a \Leftrightarrow \bigwedge_{b:(b,a) \in R} \neg b)$$

to...

Stable theory of the set A

$$th_{st}(A) = \bigwedge_{a_i \in A}(acc(a_i) \Leftrightarrow \forall a_1, \ldots, a_n, \\ (\bigwedge_{a \in A}(a \Leftrightarrow \bigwedge_{b \in A}(att(b,a) \Rightarrow \neg b)) \Rightarrow a_i))$$

Encoding the Stable Semantics (2)

From

$$\bigwedge_{a \in A} (a \Leftrightarrow \bigwedge_{b:(b,a) \in R} \neg b)$$

to...

Stable theory of the set A

$$th_{st}(A) = \bigwedge_{a_i \in A}(acc(a_i) \Leftrightarrow \forall a_1, \ldots, a_n, \\ (\bigwedge_{a \in A}(a \Leftrightarrow \bigwedge_{b \in A}(att(b,a) \Rightarrow \neg b)) \Rightarrow a_i))$$

Encoding the Stable Semantics (2)

From

$$\bigwedge_{a \in A} (a \Leftrightarrow \bigwedge_{b:(b,a) \in R} \neg b)$$

to...

Stable theory of the set A

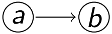
$$th_{st}(A) = \bigwedge_{a_i \in A}(acc(a_i) \Leftrightarrow \forall a_1, \ldots, a_n, \\ (\bigwedge_{a \in A}(a \Leftrightarrow \bigwedge_{b \in A}(att(b,a) \Rightarrow \neg b)) \Rightarrow a_i))$$

Decoding Tools

- $Proj_{att}(\Phi)$: projection of the models of Φ on the variables $att(x, y)$
- $arg(Mods_{att})$: generation of AFs from models projected on $att(x, y)$

Example of decoding

With $A = \{a, b\}$, the revised models could be:
$Mod(\Phi) = \{\{acc(a), \neg acc(b), \neg att(a, a), att(a, b), \neg att(b, a), \neg att(b, b)\}\}$.
So, $Proj_{att}(\Phi) = \{\{\neg att(a, a), att(a, b), \neg att(b, a), \neg att(b, b)\}\}$ and
$arg(Proj_{att}(\Phi)) = \{F\}$ with F the AF below:



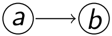
AMANDE

Decoding Tools

- $Proj_{att}(\Phi)$: projection of the models of Φ on the variables $att(x, y)$
- $arg(Mods_{att})$: generation of AFs from models projected on $att(x, y)$

Example of decoding

With $A = \{a, b\}$, the revised models could be:
$Mod(\Phi) = \{\{acc(a), \neg acc(b), \neg att(a, a), att(a, b), \neg att(b, a), \neg att(b, b)\}\}$.
So, $Proj_{att}(\Phi) = \{\{\neg att(a, a), att(a, b), \neg att(b, a), \neg att(b, b)\}\}$ and
$arg(Proj_{att}(\Phi)) = \{F\}$ with F the AF below:



Decoding Tools

- $Proj_{att}(\Phi)$: projection of the models of Φ on the variables $att(x, y)$
- $arg(Mods_{att})$: generation of AFs from models projected on $att(x, y)$

Example of decoding

With $A = \{a, b\}$, the revised models could be:
$Mod(\Phi) = \{\{acc(a), \neg acc(b), \neg att(a, a), att(a, b), \neg att(b, a), \neg att(b, b)\}\}$.
So, $Proj_{att}(\Phi) = \{\{\neg att(a, a), att(a, b), \neg att(b, a), \neg att(b, b)\}\}$ and
$arg(Proj_{att}(\Phi)) = \{F\}$ with F the AF below:

AMANDE

UNIVERSITÉ D'ARTOIS

Translation-based Revision Operator

Translation-based Revision

Let $\circ$ be a KM revision operator. For every semantics σ, every AF $F = \langle A, R \rangle$ and every formula $\varphi \in \mathcal{L}_A$, the associated *translation-based revision operator* $\star$ is defined by:

$$F \star \varphi = arg(Proj_{att}(f_\sigma(F) \circ (\varphi \wedge th_\sigma(A))))$$

AMANDE

UNIVERSITÉ D'ARTOIS

Translation-based Revision Operator

Translation-based Revision

Let $\circ$ be a KM revision operator. For every semantics σ, every AF $F = \langle A, R \rangle$ and every formula $\varphi \in \mathcal{L}_A$, the associated *translation-based revision operator* $\star$ is defined by:

$$F \star \varphi = arg(Proj_{att}(f_\sigma(F) \circ (\varphi \wedge th_\sigma(A))))$$

AMANDE

UNIVERSITÉ D'ARTOIS

Distance-based AF Revision

Let d be a distance between interpretations on $\mathcal{L}_A$. Given a formula $\psi \in \mathcal{L}_A$, the pre-order $\leq_\psi$ is defined by:

$$\omega \leq_\psi \omega' \text{ iff } d(\omega, Mod(\psi)) \leq d(\omega', Mod(\psi))$$

AMANDE

Distance-based AF Revision

Let d be a distance between interpretations on $\mathcal{L}_A$. Given a formula $\psi \in \mathcal{L}_A$, the pre-order $\leq_\psi$ is defined by:

$$\omega \leq_\psi \omega' \text{ iff } d(\omega, Mod(\psi)) \leq d(\omega', Mod(\psi))$$

The KM revision operator $\circ_d$ based on d is defined by:

$$Mod(\psi \circ_d \alpha) = \min(Mod(\alpha), \leq_\psi)$$

Distance-based AF Revision

Let d be a distance between interpretations on $\mathcal{L}_A$. Given a formula $\psi \in \mathcal{L}_A$, the pre-order $\leq_\psi$ is defined by:

$$\omega \leq_\psi \omega' \text{ iff } d(\omega, Mod(\psi)) \leq d(\omega', Mod(\psi))$$

The KM revision operator $\circ_d$ based on d is defined by:

$$Mod(\psi \circ_d \alpha) = \min(Mod(\alpha), \leq_\psi)$$

The AF revision operator $\star_d$ based on distance d is defined by:

$$F \star_d \varphi = arg(Proj_{att}(f_\sigma(F) \circ_d (\varphi \wedge th_\sigma(A))))$$

Distances and Minimal Change

Priority to Minimal Change on Acceptance Statuses

Let A be a set of arguments, and $N = |A|^2 + 1$.

$$d_H^{acc}(\omega, \omega') = \sum_{a \in A}(\omega(acc(a)) \oplus \omega'(acc(a))) + \sum_{a,b \in A}(\omega(att(a,b)) \oplus \omega'(att(a,b)))$$

Distances and Minimal Change

Priority to Minimal Change on Acceptance Statuses

Let A be a set of arguments, and $N = |A|^2 + 1$.

$$d_H^{acc}(\omega, \omega') = N \times \sum_{a \in A} (\omega(acc(a)) \oplus \omega'(acc(a))) + \sum_{a,b \in A} (\omega(att(a,b)) \oplus \omega'(att(a,b)))$$

Distances and Minimal Change

Priority to Minimal Change on Acceptance Statuses

Let A be a set of arguments, and $N = |A|^2 + 1$.

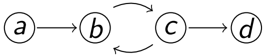
$$\begin{aligned} d_H^{acc}(\omega, \omega') = \ & N \times \sum_{a \in A} (\omega(acc(a)) \oplus \omega'(acc(a))) \\ & + \sum_{a,b \in A} (\omega(att(a,b)) \oplus \omega'(att(a,b))) \end{aligned}$$

Priority to Minimal Change on the Attack Relation

Let A be a set of arguments, and $N = |A| + 1$.

$$\begin{aligned} d_H^{att}(\omega, \omega') = \ & \sum_{a \in A} (\omega(acc(a)) \oplus \omega'(acc(a))) \\ & + N \times \sum_{a,b \in A} (\omega(att(a,b)) \oplus \omega'(att(a,b))) \end{aligned}$$

Example

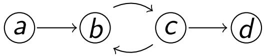
$F_1 =$ 

Accepted arguments: $\{a, c\}$

Revision by $\varphi = acc(a) \land \neg att(a, b)$ with priority to minimal change on...

AMANDE

Example

$F_1 =$ 

Accepted arguments: $\{a, c\}$

Revision by $\varphi = acc(a) \land \neg att(a, b)$ with priority to minimal change on...

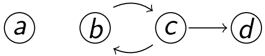
...the attack relation:

1 acceptance status change

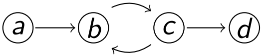
- $\{a\}$

1 attack change

- removal of (a, b)

$F_2 =$ 

Example

$F_1 =$ 

Accepted arguments: $\{a, c\}$

Revision by $\varphi = acc(a) \land \neg att(a, b)$ with priority to minimal change on...

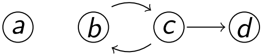
...the attack relation:

1 acceptance status change

- $\{a\}$

1 attack change

- removal of (a, b)

$F_2 =$ 

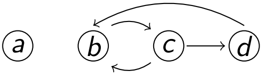
...acceptance statuses:

0 status change

- $\{a, c\}$

2 attack changes

- removal of (a, b)
- addition of (d, b)

$F_3 =$ 

Outline

Conclusion

- Characterization of other rational revision operators
- Implementations of revision operators: SAT solvers
- Other kind of change in AF (enforcement, merging of AF,...)

Conclusion

- Characterization of other rational revision operators
- Implementations of revision operators: SAT solvers
- Other kind of change in AF (enforcement, merging of AF,...)

Other interests:

- Inconsistency measures
- Logical encodings of argumentation semantics

mailly@cril.fr

AMANDE

UNIVERSITÉ D'ARTOIS

Conclusion

- Characterization of other rational revision operators
- Implementations of revision operators: SAT solvers
- Other kind of change in AF (enforcement, merging of AF,...)

Other interests:

- Inconsistency measures
- Logical encodings of argumentation semantics

mailly@cril.fr

On the Revision of Argumentation Systems: Minimal Change of Arguments Statuses In *KR'2014*

A Translation-based Approach for Revision of Argumentation Frameworks In *JELIA'2014*