An overview of argumentation frameworks for decision support

Francesca Toni

Department of Computing, Imperial College London, UK

Eighth Workshop in Decisions, Games and Logic
LSE, 17 June 2015

Outline

- Part I:
  - Argumentation frameworks for
    - non-monotonic reasoning/logics
    - (game theory)
    - classical propositional reasoning/logic and beyond
- Part II:
  - Argumentation frameworks for decision support:
    - *bipolar* argumentation frameworks and collaborative *Q&A*-based decisions
    - *value-based* argumentation frameworks and collaborative *multi-agent* decisions
    - *assumption-based* argumentation fameworks and *multi-criteria* decisions

Non-Monotonic Reasoning

A person is innocent unless proven guilty

Jo is a person, accused of some crime: should Jo be deemed innocent? **Yes**

New evidence indicates beyond reasonable doubt that Jo committed the crime: should Jo be deemed innocent now? **No**

Non-Monotonic Logics

A person is innocent unless proven guilty

Jo is a person, accused of some crime; evidence indicates beyond reasonable doubt that she committed the crime: should Jo be deemed innocent? **No, no matter which Non-Monotonic Logic**:

- Logic Programming with Negation as Failure:
  $$Inn(x) \leftarrow Acc(x), not\ Guilty(x),$$
  $$Acc(Jo) \leftarrow, \quad Guilty(Jo) \leftarrow$$
- Default Logic:
  $$D : \frac{Acc(Jo) : M\neg Guilty(Jo)}{Inn(Jo)}$$
  $$W : Acc(Jo), \quad Guilty(Jo)$$
- Non-Monotonic Modal Logic:
  $$Acc(Jo) \wedge \neg LGuilty(Jo) \rightarrow Inn(Jo),$$
  $$Acc(Jo), \quad Guilty(Jo)$$
- ...

Argumentation for NMR: intuition

- LP with NAF:
  $Inn(x) \leftarrow Acc(x), not\ Guilty(x), \quad Acc(Jo) \leftarrow, \ Guilty(Jo) \leftarrow$
  - there is an **argument** for $Inn(Jo)$ supported by $not\ Guilty(Jo)$
  - there is an objection (**attack**) against this argument, namely an argument for $Guilty(Jo)$
  - there is no objection (**attack**) against this argument
  - (the argument for) $Inn(Jo)$ is thus not "**acceptable**"
- Default Logic:
  $$D : \frac{Acc(Jo) : M\neg Guilty(Jo)}{Inn(Jo)}, \ W : Acc(Jo), Guilty(Jo)$$
  - there is an **argument** for $Inn(Jo)$ supported by $M\neg Guilty(Jo)$
  - there is an objection (**attack**) against this argument, namely an argument for $Guilty(Jo)$
  - there is no objection (**attack**) against this argument
  - (the argument for) $Inn(Jo)$ is thus not "**acceptable**"
- Non-Monotonic Modal Logic: ...

Abstract Argumentation (AA) frameworks[1]

- An AA framework is a pair $\langle Args, attacks \rangle$ where
  - $Args$ is a set (the arguments)
  - $attacks \subseteq Args \times Args$ is a binary relation over $Args$ ($(\alpha, \beta) \in attacks$ often written as "α attacks β")
- Equivalently, an AA framework is a directed graph

e.g. argument for $Inn(Jo) \longleftarrow$ argument for $Guilty(Jo)$

Given arguments and attacks between them, what is **acceptable**?

e.g. (the argument for) $Guilty(Jo)$ is,
(the argument for) $Inn(Jo)$ is not

[1] Dung 1995

Acceptability in AA frameworks

A set of arguments (*extension*) is

- *conflict-free* iff it does not attack itself
- *stable* iff it is conflict-free and attacks every argument it does not contain
- *admissible* iff it is conflict-free and attacks back each attacking argument; *preferred* iff it is maximal (wrt $\subseteq$) admissible
- *complete* iff it is admissible and contains all arguments it defends (by attacking all attacks against them); *grounded* iff it is minimal (wrt $\subseteq$) complete
- ...

e.g. given AA framework $\alpha \longleftrightarrow \beta$
$\{\alpha\}$ is stable, preferred, complete (and so is $\{\beta\}$); $\{\}$ is grounded

Arguments in AA

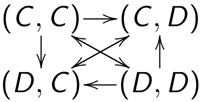
- Deductions (in some underlying logic) – (stable) extensions correspond to entailment in Non-Monotonic Logics
- Pairs of strategies in Games – stable extensions correspond to Nash equilibria
- Sets of sentences in classical Propositional Logic (PL) – a notion of acceptable arguments corresponds to classical entailment for consistent theories but does not trivialise for inconsistent theories → Argumentation Logic (AL)[2]

[2] Kakas, Mancarella, Toni 2014

AA for Game Theory

	C	D
C	1,2	0,0
D	0,0	2,1

Nash Equilibria:
(C,C) and (D,D)



Stable extensions:
$\{(C,C)\}$ and $\{(D,D)\}$

(X,Y) attacks (X',Y') iff

- $X \neq X'$ and $Y \neq Y'$, or
- $X = X'$ and $u_{a_2}(X,Y) \geq u_{a_2}(X,Y')$, or
- $Y = Y'$ and $u_{a_1}(X,Y) \geq u_{a_1}(X',Y)$

AL for PL – Preliminaries

Let T be a theory in PL and ϕ a sentence:

- a *direct derivation* for ϕ (from T) is a Natural Deduction derivation of ϕ (from T) without any application of Reductio ad Absurdum
- $T \vdash_{MRA} \phi$ denotes a direct derivation for ϕ from T

E.g. $T \vdash_{MRA} \alpha$ for $T = \{\alpha \wedge \beta\}$:

$\alpha \wedge \beta$	from T
α	$\wedge E$

but $T \nvdash_{MRA} \gamma$ for $T = \{\alpha, \neg\alpha\}$, although $T \vdash \gamma$:

$\lceil \neg\gamma$	hypothesis
α	from T
$\neg\alpha$	from T
$\bot \rfloor$	$\wedge I$
$\neg\neg\gamma$	$\neg I$ (RA)
γ	$\neg E$

- *Directly inconsistent* theory: $T \vdash_{MRA} \bot$
  Directly consistent theory: $T \nvdash_{MRA} \bot$ (but possibly $T \vdash \bot$)

AL for PL: Frameworks

Let T be a PL theory (over $\mathcal{L}$). $\langle Args^T, Att^T, Def^T \rangle$ consists of

- $Args^T = \{T \cup \Sigma | \Sigma \subseteq \mathcal{L}\}$
- given $a, b \in Args^T$, with $a = T \cup \Delta$, $b = T \cup \Gamma$, such that $\Delta \neq \{\}$, $(b, a) \in Att^T$ iff $a \cup b \vdash_{MRA} \bot$
  b attacks a stands for $(b, a) \in Att^T$
- given $a, d \in Args^T$, with $a = T \cup \Delta$, $(d, a) \in Def^T$ iff
  1. $d = T \cup \{\neg\phi\}$ ($d = T \cup \{\phi\}$) for some sentence $\phi \in \Delta$ (respectively $\neg\phi \in \Delta$), or
  2. $d = T \cup \{\}$ and $a \vdash_{MRA} \bot$

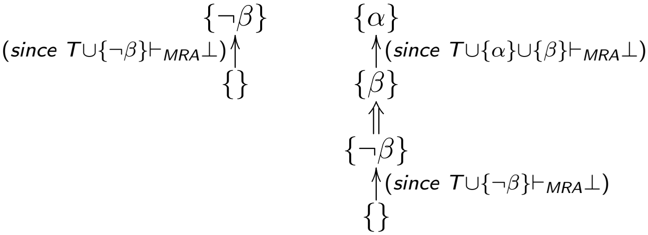
  d defends against a stands for $(d, a) \in Def^T$

AL for PL: Acceptability

For T *directly consistent*, $a, a_0 \in Args^T$, $NACC^T(a, a_0)$ iff

- $a \not\subseteq a_0$ and
- there exists $b \in Args^T$ s.t. b attacks a and
  - $b \subseteq a_0 \cup a$, or
  - for all $d \in Args^T$ s.t. d defends against b: $NACC^T(d, a_0 \cup a)$

e.g. $T = \{\alpha \wedge \beta \rightarrow \bot, \neg\beta \rightarrow \bot\}$:



i.e. $NACC^T(\{\neg\beta\}, \{\})$ $\quad$ $NACC^T(\{\alpha\}, \{\})$

Argumentation Logic

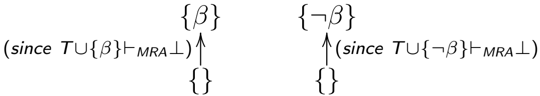
- $ACC^T(\{\phi\},\{\}) = \text{not } NACC^T(\{\phi\},\{\})$
- for T directly consistent:
  ϕ is *AL-entailed by T* (denoted $T \models_{AL} \phi$) iff
  $ACC^T(\{\phi\},\{\})$ and $NACC^T(\{\neg\phi\},\{\})$

AL corresponds to classical PL for classically consistent theories:

- for T *classically consistent* (expressed using only $\neg$ and $\wedge$):
  $T \models_{AL} \phi$ iff $T \vdash \phi$

but does not trivialise:

- $T = \{\neg(\alpha \wedge \beta), \neg(\gamma \wedge \neg\beta), \alpha, \gamma\}$ is *classically inconsistent*, but
  $T \not\models_{AL} \beta$ and $T \not\models_{AL} \neg\beta$
  since both $NACC^T(\{\beta\},\{\})$ and $NACC^T(\{\neg\beta\},\{\})$:



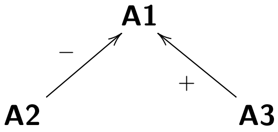
Summary (up until now)

- Part I:
  - Abstract Argumentation (AA) frameworks for
    - non-monotonic reasoning/logics
    - (game theory)
  - A variant of an instance of AA frameworks – called Argumentation Logic (AL) – for classical propositional reasoning/logic and beyond
- Part II:
  - Argumentation frameworks for decision support:
    - *bipolar* argumentation frameworks and collaborative *Q&A*-based decisions
    - *value-based* argumentation frameworks and collaborative *multi-agent* decisions
    - *assumption-based* argumentation fameworks and *multi-criteria* decisions

Bipolar Argumentation (BA) Frameworks[3]

- A BA framework is a triple $\langle Args, attacks, supports \rangle$ where
  - $\langle Args, attacks \rangle$ is an AA framework and
  - $supports \subseteq Args \times Args$ is a binary relation over $Args$ ($(\alpha, \beta) \in supports$ often written as "α supports β")

John: *I think we should go and see the new Avengers; the first one was great!* (**A1**)
Joe: Please spare me! It's just going to be another big Hollywood production that goes for explosions instead of plot. (**A2**)
Jane: *I loved the first one, as well, so I think we should see it!* (**A3**)



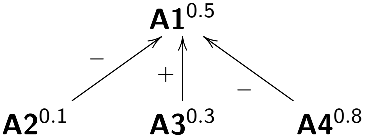
[3] C. Cayrol and M. Lagasquie-Schiex 2005

From Acceptability to Strength

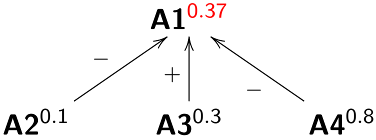
- $\{$**A1**$\}$ is not acceptable (under any notion of acceptability) both on the left and on the right
- but $\{$**A1**$\}$ is less "strong" on the right!

Strength of arguments[4]

- Consider BA frameworks in the restricted form of *trees*
- Let arguments have a *base score* (a real number in [0,1]) e.g. derived from *votes*
- Strength of arguments computed from leaves to root, e.g.



BA with base scores



BA with computed strengths

strength of **A1** =

(strength after attacks + strength after support)/2 =

(0.09+0.65)/2 = 0.74/2 = 0.37

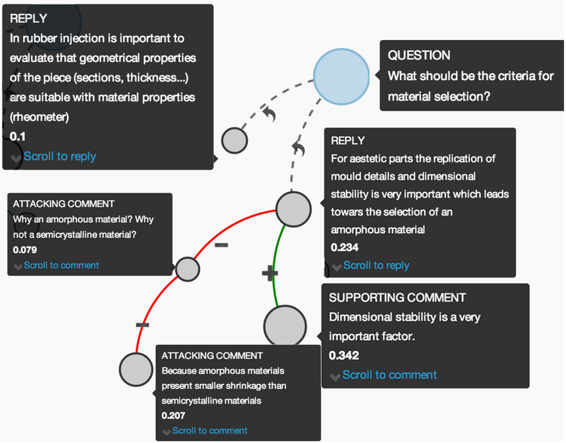
strength after attacks= (0.5 - 0.5*0.1) - (0.5 - 0.5*0.1)*0.8 = 0.09

strength after support= 0.5+(1-0.5)*0.3 = 0.65

[4] Baroni et al 2015

Collaborative Q&A-based decisions - 1

Quaestio-it (`www.quaestio-it.com`) and Desmold

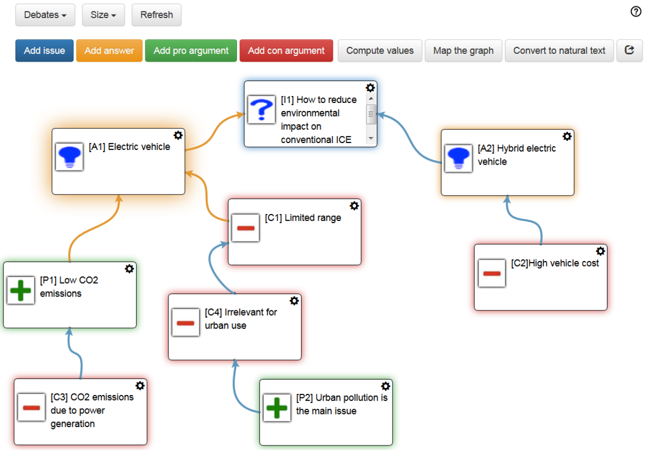


(positive, negative and spam) votes

Collaborative Q&A-based decisions - 2

Arg&Dec (www.arganddec.com)

Link to matrix-based decisions, natural language explanations



Value-based AA[5]: a simple example

- Consider an AA framework $a \leftrightarrow b$ where

  a: Let's have dinner at home today

  b: Let's have dinner in a restaurant today

  - Grounded extension $\{\}$; two preferred extensions: $\{a\}$ and $\{b\}$

- Assume arguments *promote values*

  $v1$: Money-saving, a promotes $v1$

  $v2$: Time-saving, b promotes $v2$

  - $v1 > v2$: $\{a\}$ should be grounded/the only preferred extension;
  - $v2 > v1$: $\{b\}$ should be grounded/the only preferred extension

- Value-based AA uses preferences over values promoted by arguments to obtain a *simplified AA framework*:

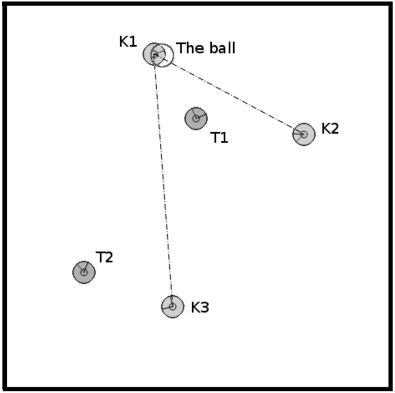
  if $v1 > v2$ then $a \rightarrow b$

  if $v2 > v1$ then $a \leftarrow b$

[5] Bench-Capon 2003

Collaborative Multi-Agent Decisions

Example: Takeaway in Robocup



Agents T1 and T2 (the “takers”) need to collaborate to get the ball from K1, K2, K3 (the “keepers”)

Value-based AA for Takeaway[6]

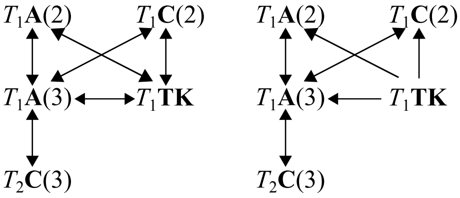
- Values:
  1. **VT**: Prevent the ball being held by the keepers;
  2. **VO**: Prevent the ball being passed to an 'open' keeper;
  3. **VF**: Prevent the ball being passed to a 'far' keeper;
  4. **VA**: Ensure that each pass can be quickly intercepted;
  5. **VC**: Ensure that, after each pass, the ball can be tackled.
- Arguments promote values:
  1. $T_i\mathbf{TK}$ promotes **VT**
  2. $T_i\mathbf{O}(p)$ promotes **VO**
  3. $T_i\mathbf{F}(p)$ promotes **VF**
  4. $T_i\mathbf{A}(p)$ promotes **VA**
  5. $T_i\mathbf{C}(p)$ promotes **VC**
- Ranking over values:
  $\mathbf{VT} > \mathbf{VA} = \mathbf{VC} > \mathbf{VO} > \mathbf{VF}$

[6] Gao, Toni 2014

Value-based AA for Takeaway: an example

T_i**TK** promotes **VT**, T_i**O**(p) promotes **VO**, T_i**F**(p) promotes **VF**, T_i**A**(p) promotes **VA**, T_i**C**(p) promotes **VC**

VT > **VA** = **VC** > **VO** > **VF**



- grounded extension of the simplified AA framework (right) obtained from the Value-based AA (left): $\{T_1$**TK**, T_2**C**(3)$\}$
- T_1 should tackle the ball, T_2 should mark K_3

Assumption-Based Argumentation (ABA)[7]

- A form of *structured argumentation*: differently from AA, arguments and attacks are *not* primitive notions
- An instance of AA
- Admits AA as an instance

[7] Bondarenko et al 1997

ABA frameworks

An ABA framework is a tuple $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \mathcal{C} \rangle$ where

- $\langle \mathcal{L}, \mathcal{R} \rangle$ is a deductive system with language $\mathcal{L}$ and rules $\mathcal{R}$,
- $\mathcal{A} \subseteq \mathcal{L}$ are assumptions,
- $\overline{}$ is a total mapping from $\mathcal{A}$ into $\mathcal{L}$, $\overline{\alpha}$ is the contrary of α.

Arguments are deductions (wrt $\langle \mathcal{L}, \mathcal{R} \rangle$) of claims supported by sets of assumptions.

Attacks are directed at the assumptions in the support of arguments – by deriving their contrary.

e.g. for $\mathcal{R} = \{s \leftarrow b, t \leftarrow c\}$, $\mathcal{A} = \{a, b, c\}$, $\overline{a} = s, \overline{b} = t, \overline{c} = u$:

$$\{a\} \vdash a \longleftarrow \{b\} \vdash s \longleftarrow \{c\} \vdash t$$

ABA semantics

- stable, preferred, grounded etc sets of arguments – as in AA
- stable, preferred, grounded etc sets of assumptions

The two views (argument view and assumption view) correspond

e.g. for $\mathcal{R} = \{s \leftarrow b, t \leftarrow c\}$, $\mathcal{A} = \{a, b, c\}$, $\overline{a} = s, \overline{b} = t, \overline{c} = u$:
$\{b\}$ attacks $\{a\}$, $\{c\}$ attacks $\{b\}$

- $\{c, a\}$ is a stable etc set of assumptions
- the set of all arguments supported by subsets of $\{c, a\}$ is a stable etc extension

Multi-attributes decision making

Example: Attending a workshop at Imperial: accommodation?

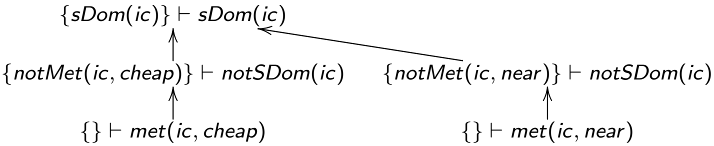
Decisions: jh, ic, ritz
Attributes: £50, £70, £200, sk, pic
Goals: cheap, near

	£50	£70	£200	sk	pic
jh		✓		✓	
ic	✓			✓	
ritz			✓		✓

	£50	£70	£200	sk	pic
cheap	✓				
near				✓	

ic is **strongly dominant**: it meets all goals

From Multi-Attribute Decision Making to ABA[8]



- d is strongly dominant iff $\{sDom(d)\} \vdash sDom(d)$ belongs to an admissible set

[8]Fan, Toni 2013, EPSRC TRaDAr project

Conclusions

- Various types of argumentation fameworks
- supporting various types of decision-making
- and providing explanations for decisions
- while dealing with potentially incomplete and inconsistent information

Bibliography

- L. Carstens, X. Fan, Y. Gao and F. Toni, An overview of argumentation frameworks for decision support, The Fourth IJCAI International Workshop on Graph Structures for Knowledge Representation and Reasoning (GKR 2015). To appear.
- M. Aurisicchio, P. Baroni, D. Pellegrini, and F. Toni, Comparing and integrating argumentation-based with matrix-based decision support in Arg&Dec, Third International Workshop on Theory and Applications of Formal Argumentation (TAFA-15). To appear
- A. Kakas, F. Toni and P. Mancarella, Argumentation Logic, Fifth International Conference on Computational Models of Argument, COMMA 2014
- X. Fan and F. Toni, Decision making with assumption-based argumentation, Theories and Applications of Formal Argumentation - Second International Workshop, TAFA 2013, volume 8306 of LNAI, Springer, 2013
- Bibliography of these papers