Reasoning about Preferences in Structured Extended Argumentation Frameworks

Sanjay Modgil[1]

Henry Prakken[2]

1 Department of Informatics, King's College London

2 Department of Information and Computing Sciences, Utrecht University & Faculty of Law, University of Groningen

Outline of Talk

- Background

  1) ASPIC+ : Structuring Dung's Abstract Argumentation Theory to identify a range of possible instantiations $\Rightarrow$ extra expressivity allows for study of rationality postulates

  2) Extended Abstract Argumentation Theory : incorporates meta-argumentation *about* preferences applied to arguments

- Structuring Extended Argumentation Theory

- Rationality postulates satisfied by structured extended argumentation theories under assumptions

ASPIC+ [1]

- ASPIC+ builds on ASPIC: 1) defines a more general class of instantiations of Dung AFs; 2) Shows that rationality postulates satisfied when preferences over arguments accounted for

- Tree structured arguments (Vreeswijk) built from

  - ***defeasible*** *and* ***strict*** *inference rules* $\varphi_1 \ldots \varphi_n \Rightarrow \varphi$ *and* $\varphi_1 \ldots \varphi_n \rightarrow \varphi$ *where each* φ_i *is a wff in some logical language* $\mathcal{L}$

  - a knowledge base $\mathcal{K}$ *of premises that are wff in* $\mathcal{L}$

    1. Ordinary premises

    2. Axiom premises

    3. Assumption premises

1 H. Prakken. An abstract framework for argumentation with structured arguments. To appear in: *Argument and Computation*, **1**, (2010).

ASPIC+

- Negation: generalised to arbitrary contrary relation over wff of $\mathcal{L}$ (cf. ABA: need not be symmetric)
- Contrary based attacks defined :
  - on ordinary or assumption premises (undermining),
  - on conclusion of **defeasible** inference rule (rebutting),
  - on defeasible inference rule itself (undercutting)
- *A* defeats *B* iff for some sub-argument B′ of B, *A's* attack on *B′* *succeeds*
  - Some types of attack always succeed (e.g. undercut, undermining attack on an assumption premise)
  - Some succeed only if not $A <_a B'$, where $<_a$ is an argument ordering defined by given orderings on defeasible rules and non-axiom premises

Rationality Postulates and ASPIC+

- Arguments and defeats instantiate a Dung framework
- Closure and Consistency rationality postulates* shown to be satisfied under certain assumptions (e.g., strict rules closed under transposition or contraposition)

- Significance of ASPIC+ is that it structures Dung frameworks, identifying a a *range* of possible instantiations that satisfy rationality postulates, e.g.,
  - argument schemes can be represented
  - assumption based argumentation shown to be a special case
  - more recently deductive argumentation (e.g., Besnard and Hunter) shown to be a special case
  - argument orderings assumed to be defined by weakest or last link principles

* M. Caminada and L. Amgoud. On the evaluation of argumentation formalisms. *Artificial Intelligence*,**171**, 286–310, (2007).

Extended Argumentation Theory [1]

- Extended Argumentation Framework (EAF) = (*Args*, *Attack*, *PrefAtt*)
- $PrefAtt \subseteq (Args \times Attack)$
  if $(z, (x,y)) \in PrefAtt$ then z expresses that y is preferred to x
- Ensuring satisfaction of consistency postulates motivated two features of *abstract* theory :

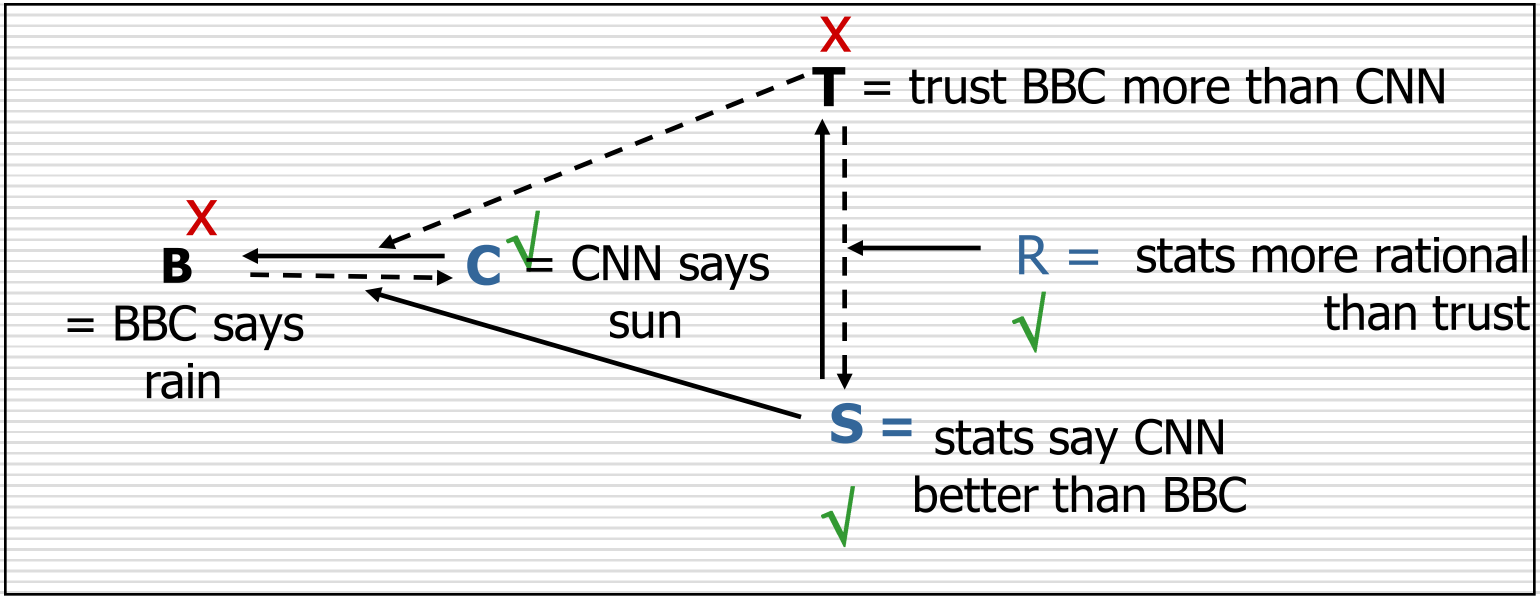
1) if $(z, (x,y)), (z', (y,x)) \in PrefAtt$ then $(z,z'), (z',z) \in Attack$

2) $S \subseteq Args$ is conflict free if it *contains no symmetrically attacking arguments*, and if $x,y \in S$, $(x,y) \in Attack$ then $z \in S$, $(z, (x,y)) \in PrefAtt$

1 S. Modgil. Reasoning about preferences in argumentation frameworks. *Artificial Intelligence*, **173**, 901–934, (2009).

Extended Argumentation Theory

- Modified def. of acceptability defined for EAFs
- Extensions of EAFs under Dung semantics defined



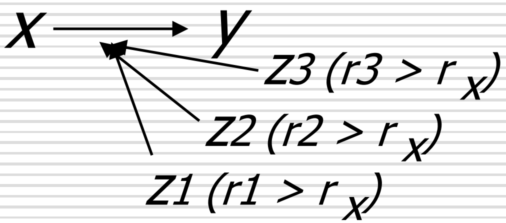
- { C, S, R } is single grounded / preferred extension

Structuring EAFs

- EAFs subsume and extend preference and value-based argumentation to accommodate argumentation based reasoning about possibly contradictory preferences/values
- Adopt Dung's level of abstraction - hence provide for instantiation by logics facilitating reasoning about priorities (over names of object level formulae)
- However, as with Dung AFs, level of abstraction precludes identification of appropriate (in terms of satisfying rationality postulates) instantiations
- $\Rightarrow$ apply ASPIC+ methodology to EAFs ...

Re-defining the Extended Theory with *Collective* Pref-Attacks

- EAFs with collective pref attacks = (*Args*, *Attack*, *PrefAtt*)
- *PrefAtt* $\subseteq (2^{Args}/\emptyset) \times Attack$)



- Features of *abstract* theory designed to ensure that consistency postulates satisfied:

1) if ~~$(z , (x,y)), (z', (y,x)) \in PrefAtt$ then $(z,z'), (z',z) \in Attack$~~

2) S $\subseteq$ *Args* is conflict free ~~if it *contains no symmetrically attacking arguments*, and~~ if $x,y \in S$, $(x,y) \in Attack$ then $z \in S$, $(z , (x,y)) \in PrefAtt$

- ***All properties of Extended Argumentation Theory preserved***

Structuring EAFCs

- Tree structured arguments built from *defeasible and strict inference rules* and a knowledge base of ordinary, axiom and assumption premises (as for ASPIC+), but now ***no given argument ordering***

- Contrary relation defined over language, and contrary based attack relation defined as for ASPIC+

- Partial function $\mathcal{P}$ maps sets of arguments to pairs (Y,X)
  e.g., $(Y,X) \in \mathcal{P}(\{Z1, Z2, Z3\})$ means *Z1, Z2, Z3* collectively conclude that *Y* is preferred to *X*

Structuring EAFCs

- (*Args*,*Attack*,*PrefAtt*) where *Args* and *Attack* defined as for ASPIC+
- $(\Phi, (x,y)) \in$ *PrefAtt* if $(x,y) \in$ *Attack*, and:

  $\forall$ sub-argument y′ of y, s.t. x rebuts or undermines y′:
  - (x,y′) does not succeed independently of preferences
  - $\exists \Phi' \subseteq \Phi$ such that $(y',x) \in \mathcal{P}(\Phi')$
  - Φ is set inclusion minimal

Example

X = [op1 : ¬a]

Z = [ap1 : c , d1 : c ⇒ op1 > op2]

Y = [op2 : b , s1 : b → a]

(*X*,*Y*) ∈ $\mathcal{P}$({*Z*}) by weakest and last link principles

Q = [op3 : e , d2 : e ⇒ ¬c]

Rationality Postulates

- **Theorem Sub-argument Closure**: Let E be a complete extension of a structured EAFC. For any A $\in$ E, if A' is a sub-argument of A then A' $\in$ E

- **Theorem Closure under Strict Rules**: Let E be a complete extension of a structured EAFC. Then {Conc(A)| A $\in$ E} = Cl_{Rs} ({Conc(A) | A $\in$ E}).

- Satisfaction of consistency postulates under assumptions for ASPIC+ ***and*** "preference assumptions" on $\mathcal{P}$ that have been shown to be satisfied by *weakest* and *last link* principles:

Rationality Postulates

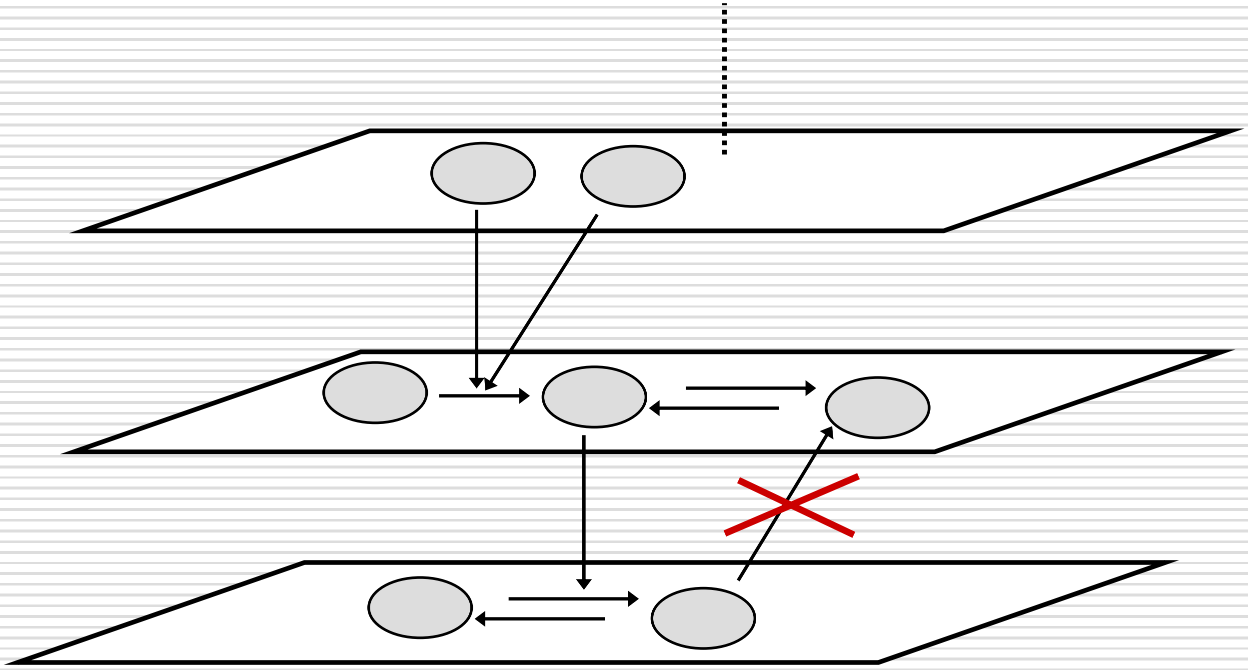
Not in COMMA paper, but more recently proved that:

- **Theorem Direct Consistency**: Let E be the grounded extension of a structured EAFC. Then {Conc(A) | A $\in$ E} is consistent
- **Theorem Indirect Consistency** : Let E be the grounded extension of a structured EAFC. Then Cl_{Rs}({Conc(A) | A $\in$ E}) is consistent.

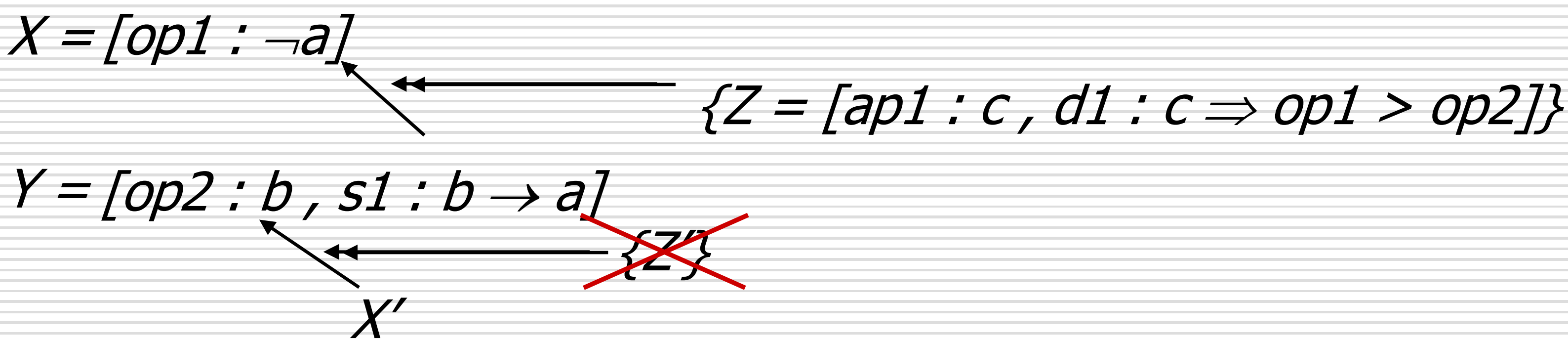
In COMMA paper the consistency theorems are shown for *any semantics* subsumed under the complete semantics, for *hierarchical* EAFCs

Hierarchical *EAF*s

- Hierarchical *EAF*s restrict interactions between the levels - shown to suffice for applications of extended argumentation to agent, normative, and legal reasoning



Example



- $E = \{X, Y, Z\}$ is conflict free and complete but would seem to violate direct consistency !
- But X can be extended with transposition s1′ : $\neg a \rightarrow \neg b$, obtaining $X' =$ [$op1 : \neg a$, s1′ : $\neg a \rightarrow \neg b$] which must be in E
- If $\mathcal{P}$ satisfies preference assumptions (as shown for *weakest* and *last link*), *there cannot be a* $Z' \in E$ *s.t.* $(\{Z'\},(X',Y))$ making E’ = $\{X, Y, Z, X', Z'\}$ conflict free and complete
- => there cannot be such an E = $\{X, Y, Z\}$ violating direct consistency

Conclusions

- Structured EAFCs provide for instantiation by a *range* of logics, **while ensuring satisfaction of rationality postulates**

  e.g., existing logics that encode object level reasoning about priorities (e.g., Prakken and Sartor's logic programming with defeasible priorities)

- Currently extending special cases of ASPIC+ to accommodate argumentation reasoning about preferences over arguments

  - e.g., extending deductive argumentation (Besnard and Hunter) with classical logic arguments that claim preferences over other classical logic arguments