A Canonical Semantics for Structured Argumentation with Priorities

Phan Minh Dung

Asian Institute of Technologies, Thailand

September 13, 2016

Outlines

1. Motivating Example
2. Regular Properties
3. Regular Attack Relations Assignments
4. Upper Semi-Lattice of Regular Attack Relation Assignments
5. Unique Canonical Semantics as Greatest Regular Attack Relation Assignments

A Sherlock Holmes Investigation (From Dung: AIJ2016)

Sherlock Holmes is investigating a case involving three persons P_1, P_2 and S together with the dead body of a big man.

1. The knowledge that one of the persons is the murderer is represented by three strict rules:

$$r_1: \; Inno(P_1), Inno(S) \rightarrow \neg Inno(P_2)$$

$$r_2: \; Inno(P_2), Inno(S) \rightarrow \neg Inno(P_1)$$

$$r_3: \; Inno(P_1), Inno(P_2) \rightarrow \neg Inno(S)$$

2. S is a small child who cannot kill a big man. This fact is captured in the base of evidence $BE = \{Inno(S)\}$.

A Sherlock Holmes Investigation

The legal principle that people are considered innocent until proven otherwise could be represented in two ways:

- By three defeasible rules

$$d_1 : \ \Rightarrow Inno(P_1) \qquad d_2 : \ \Rightarrow Inno(P_2) \qquad d : \ \Rightarrow Inno(S)$$

- By two defeasible rules

$$d_1 : \ \Rightarrow Inno(P_1) \qquad d_2 : \ \Rightarrow Inno(P_2)$$

as S is innocent, and hence the defeasible rule $d : \ \Rightarrow Inno(S)$ is intuitively redundant.

A Sherlock Holmes Investigation

- After digging around, it becomes clear to Holmes that P_1 has a strong motive to kill the victim while there is nothing connecting P_2 to the dead man.
- Holmes hence focuses his investigation on P_1.

This knowledge is represented by a preference

$$d_1 : \Rightarrow Inno(P_1) \quad \prec \quad d_2 : \Rightarrow Inno(P_2)$$

stating that *Holmes gives higher priority (in his investigation) to the scenario in which P_2 is innocent than to the other one.*

A Sherlock Holmes Investigation

- KB_0 :
  - $r_1 : \ Inno(P_1), Inno(S) \rightarrow \neg Inno(P_2)$
    $r_2 : \ Inno(P_2), Inno(S) \rightarrow \neg Inno(P_1)$
    $r_3 : \ Inno(P_1), Inno(P_2) \rightarrow \neg Inno(S)$
  - $d_1 : \ \Rightarrow Inno(P_1)$ $\qquad d_2 : \ \Rightarrow Inno(P_2)$
  - $d_1 \prec d_2$
  - $BE = \{Inno(S)\}$.
- KB_1 is obtained by adding to KB_0 the redundant default $d : \Rightarrow Inno(S)$
- Due to the fact that S is innocent, we expect that default d will have no impact on the belief sets of the knowledge base KB_1.
- *Both KB_1 and KB_0 are expected to have identical belief sets*, concluding $\neg Inno(P_1), Inno(P_2)$

- Surprisingly, KB_0, KB_1 have different belief sets wrt the semantics based on different well-known attack relations proposed in the literature.
- Some well-known proposal also allows an stable extension justifying:

$$Inno(P_1), \neg Inno(P_2)$$

Counter-intuitive to our commonsense as P_1 has a strong motive to kill the victim while there is nothing connecting P_2 to the dead man.

What went wrong here ?

We need principles and guidelines for determining the attack relations between arguments.

Knowledge Bases

1. A **rule-based system** is a triple $\mathcal{R} = (RS, RD, \preceq)$ where
    - RS is a set of strict rules,
    - RD is a set of defeasible rules, ,
    - $\preceq$ is a transitive relation over RD representing the preferences between defeasible rules.
2. A **knowledge base** is pair $K = (\mathcal{R}, BE)$
    - $\mathcal{R}$: a rule-based system,
    - BE: a *base of evidence*, representing unchallenged observations, facts ect..

Priorities between rules are very common in practical reasoning:

- Federal rules have higher priorities than state rules or
- In a family, rules made by the father have higher priorities than those made by big brother.

Brewka 1989, Gelfond and Son 1997, Brewka and Eiter 1999, Delgrande and Schaub and Tompit 2003, Modgil and Prakken 2012, 2013.

Attack Relation: A Minimal Interpretation of Priorities

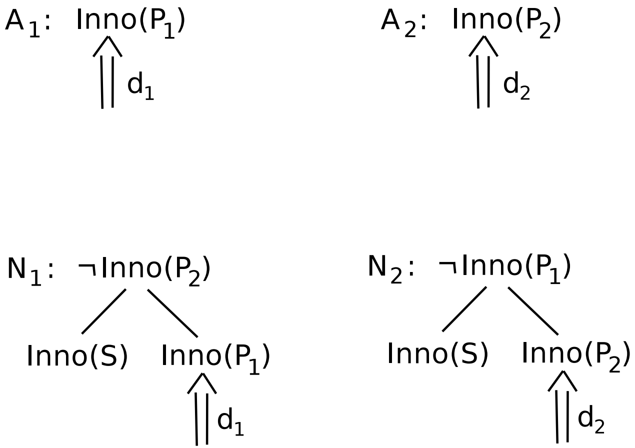
$d_0 \prec d_1$: In situations when both are applicable but applying both d_0, d_1 is not possible, d_1 should be preferred.

Effective Rebut: Given two arguments A_1, A_2 s.t.

- each A_i, $i = 1, 2$, contains exactly one defeasible rule d_i,
- A_2 contradicts A_1 and last rule in A_1 is defeasible.

A_2 attacks A_1 iff $d_2 \not\prec d_1$.

Attack Relation: A Minimal Interpretation of Priorities



N_2 attacks A_1 because $d_2 \not\prec d_1$.

N_1 does not attacks A_2 because $d_1 \prec d_2$.

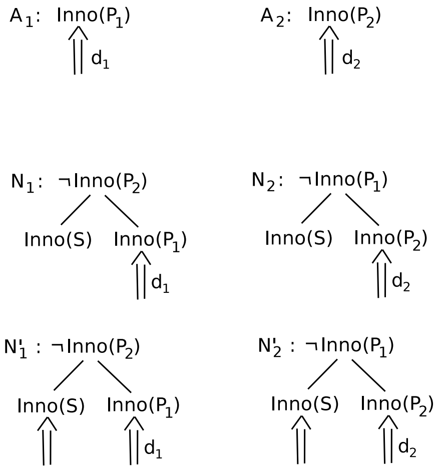
Attack Relations: Subargument Structure

If A attacks a subargument of B
then A attacks B.

Attack Relation: Weakening of Arguments

An argument is weakened when some of its facts are replaced by defeasible beliefs.

N_1', N_2' are weakened versions of N_1, N_2.



Attack Relations: Attack Monotonicity

1. If A attacks B and D is a weakening of B then A attacks D
2. If C attacks B and C is a weakening of A then A attacks B.

Attack Relations: Inconsistency Resolving

An attack relation satisfies the **inconsistency-resolving property** iff for each finite set of arguments S, if S is inconsistent then S is attacked by some argument generated by S.

Attack Relations: Link-Orientation

Intuition: In real world conversation, if you claim that my argument is wrong, I would naturally ask which part of my argument is wrong.

Regular Attack Relation Assignments: Context-Independence

$\mathcal{R} = (RS, RD, \preceq)$: a rule-based system.

$\mathcal{C}_{\mathcal{R}}$: The class of all consistent knowledge bases of the form $(\mathcal{R}, BE)$.

An **attack relation assignment atts** for a rule-based system $\mathcal{R}$ is a *function* assigning to each knowledge base $K \in \mathcal{C}_{\mathcal{R}}$ an attack relation $atts(K) \subseteq AR_K \times AR_K$.

Regular Attack Relation Assignments: Context-Independence

$\mathcal{C}_{\mathcal{R}}$: The class of all consistent knowledge bases of the form $(\mathcal{R}, BE)$.

An **attack relation assignment atts** for a rule-based system $\mathcal{R}$ satisfies the property of *context-independence* iff

- for any two knowledge bases $K, K' \in \mathcal{C}_{\mathcal{R}}$ and
- for any A, B from $AR_K \cap AR_{K'}$,
- it holds: $(A, B) \in atts(K)$ iff $(A, B) \in atts(K')$

Regular Attack Relation Assignments

$\mathcal{C}_{\mathcal{R}}$: The class of all consistent knowledge bases of the form $(\mathcal{R}, BE)$.

An **attack relation assignment atts** for a rule-based system $\mathcal{R}$ is **regular** iff

- atts satisfies context-independence property, and
- For each $K \in \mathcal{C}_{\mathcal{R}}$: atts(K) satisfies the properties of Effective Rebuts, Attack Monotonicity, Subargument Structure, Inconsistency-Resolving, Link-Orientation.

Semi-Lattice

A partial order $\leq$ on a set S is a semilattice iff each subset of S has a supremum wrt $\leq$.

Every semilattice S has an unique greatest element denoted by $\sqcup S$.

Semi-Lattice of REgular Attack Relation Assignments

- $\mathcal{A}$: Non-empty set of attack relation assignments.
- Define $\sqcup\mathcal{A}$ by: $\forall K \in \mathcal{C}_{\mathcal{R}}$:

  $$(\sqcup\mathcal{A})(K) = \bigcup\{\, \mathit{atts}(K) \mid \mathit{atts} \in \mathcal{A} \,\}$$

- If all attack relation assignments in $\mathcal{A}$ are regular then $\sqcup\mathcal{A}$ is also regular.

Canonical Attack Relation Assignment

- The set of all regular attack relation assignments is a semilattice wrt $\subseteq$.
- The **canonical attack relation assignment** of $\mathcal{R}$ is the supremum of all regular attack relation assignments.

Why should the canonical attack relation assignment be viewed as representing the semantics of knowledge bases with priorities ?

Minimal Removal Intuition

Purpose of introducing priorities between defeasible rules:

Remove certain undesired attacks while keeping the set of removed attacks to a minimum.

Minimal Removal Intuition

$$\begin{array}{cccc} A:\ a & A_1 : \neg a & B:\ b & B_1 : \neg b \\ \Uparrow d_0 & \Uparrow d_1 & \Uparrow d_2 & \Uparrow d_3 \end{array}$$

Figure: Minimal Removal

Suppose $d_3 \prec d_2$.

The attack of B_1 against B should be removed.

But nothing is said about the other attacks.
Hence they should be kept, i.e. the attacks that should be removed should be kept to a minimum.

Minimal Removal Intuition

$\forall K \in \mathcal{C}_{\mathcal{R}}$

$Batts(K) = \{(A, B) \mid A \text{ undercuts or rebuts B}\}$

Let atts: regular attack relation assignment.

It holds: $atts \subseteq Batts$.

The set $Batts(K) \setminus atts(K)$ could be viewed as the set of attacks removed from $Batts(K)$ due to the priorities between defeasible rules.

Minimal Removal Intuition

Combining the "minimal-removal intuition" with the concept of regular attack relation assignment:

The semantics of $\mathcal{R}$ should be captured by regular atts:

such that $Batts(K) \setminus atts(K)$ is minimal, or equivalently the set $atts(K)$ is maximal.

Canonical attack relation is such unique maximal attack relation assignment!

Many thanks, Francesca and Everybody here.

I apologize for not being able to come.