Argumentation framework for multi-attribute decision-making*

Maxime Morge†    Paolo Mancarella†

†Dipartimento di Informatica
Università di Pisa

Workshop on Computational Model of Law

*is supported by the Sixth Framework IST programme of the EC, under the 035200 ARGUGRID project.
The Fox and the Hedgehog [Tetlock 06]

Corporate executives say:
- “it is annoying to listen to someone who cannot seem to make up his or her mind”
- “the most common error in decision-making is to abandon good ideas too quickly”

_Hedgehogs know one big thing, have intuitions, and never surrender._

Judges say:
- “when considering most conflicts, I can usually see how both sides could be right”
- “I prefer interacting with people whose opinions are very different from my own”

_Foxes knows many little things, interact, and change their mind._
Outline

Motivation

Abstract decision structure

Concrete data structures

Argumentation framework
  Arguments
  Relations between arguments
  Semantics

Conclusions
Influence diagram (Id) for decision analysis [Clemen. 06]
Id for criminal sentencing [Berman & Hafner. 89]
Id for criminal sentencing [Berman & Hafner. 89]

Abstract decision structure

Recommended sentence \( (g_0) \)

\[ g_2 \prec g_1 \]

\[ g_5 \prec g_4 \prec g_3 \prec g_4 + g_5 \]

Crime \( (g_1) \)

Defendant \( (g_2) \)

\[ g_7 \prec g_6 \]

Circum. \( (g_3) \)

Victimes \( (g_4) \)

Offenses \( (g_5) \)

Extenuating circum. \( (g_6) \)

Prior crimes \( (g_7) \)

\[ a_1 \prec a_2 \]

\[ a_2 \prec a_1 \]

\[ a_2 ? a_1 \]

alcohol or drug?

\[ \neg alcohol \prec alcohol \]

driving?

\[ \text{driving} \]

probation\( (a_1) \) or jail\( (a_2) \)
The vocabulary

Definition

A decision making framework $\langle \mathcal{L}, \mathcal{I}, \mathcal{T}, \prec \rangle$ is defined s.a.

- $\mathcal{L}$ is a logic language with
  - a set of goals $(g_0, g_1, g_2)$,
  - a decision $(D(a_1), D(a_2), \ldots)$,
  - a set of beliefs $(b_1, b_2, \ldots)$.

- $\mathcal{I}$ is an incompatibility relation amongst sentences in $\mathcal{L}$
  - $\mathcal{I} (b_1, \neg b_1)$
  - $\mathcal{I} (D(a_1), D(a_2), \ldots)$

- $\mathcal{T}$ is a theory, the set of statements in $\mathcal{L}$

- $\prec$ is a priority relation over $\mathcal{T}$
The statements

Definition ([Prakken & Sartor 97])

A theory $T$ is an extended logic program, i.e. a finite set of rules:

$$R : L_0 \leftarrow L_1, \ldots, L_j, \text{not } L_k, \ldots, \text{not } L_m$$

head($R$) = $L_0$.
body($R$) = \{$L_1, \ldots, \text{not } L_m$\}.

The theory compiles:

- goal rules such as $R_{12}^\alpha : g_0 \leftarrow g_1, g_2$
- epistemic rules such as $R_{12}^\beta : b_0 \leftarrow b_1, \neg b_2$
- recommending rules such as $R_1^\gamma : D(a_1) \leftarrow b_0$
- decision rules such as $R_{11}^\delta : g_1 \leftarrow D(a_1), b_0$
Preferences, uncertainty, and credibility

Definition (Priority)

The priority $\prec$ is a (partial or total) preorder on $\mathcal{T}$, i.e. a relation which is reflexive and transitive.

Different priorities for different rules:

- the priority over goal rules comes from their levels of preference, eg $R_1^\alpha : g_0 \leftarrow g_1$ has priority over $R_2^\alpha : g_0 \leftarrow g_2$
- the priority over epistemic rules comes from their levels of certainty, eg $F_1^\beta : \text{alcohol} \leftarrow$ has priority over $F_2^\beta : \neg \text{alcohol} \leftarrow$
- the priority over decision rules come from their levels of credibility, eg $R_{51}^\delta : g_5 \leftarrow D(a_1)$ has priority over $R_{52}^\delta : g_5 \leftarrow D(a_2)$
A walk through the example

\[ g_2 \prec g_1 \]

The goal theory

- \( R_{12}^\alpha : g_0 \leftarrow g_1, g_2 \)
- \( R_{345}^\alpha : g_1 \leftarrow g_3, g_4, g_5 \)
- \( R_{67}^\alpha : g_2 \leftarrow g_6, g_7 \)
- \( R_{45}^\alpha : g_1 \leftarrow g_4, g_5 \)
- \( R_1^\alpha : g_0 \leftarrow g_1 \)
- \( R_3^\alpha : g_1 \leftarrow g_3 \)
- \( R_6^\alpha : g_2 \leftarrow g_6 \)
- \( R_2^\alpha : g_0 \leftarrow g_2 \)
- \( R_4^\alpha : g_1 \leftarrow g_4 \)
- \( R_7^\alpha : g_2 \leftarrow g_7 \)
- \( R_5^\alpha : g_1 \leftarrow g_5 \)
A walk through the example

The epistemic theory

\[ F_1^\beta : \text{alcohol} \leftarrow \]
\[ F_2^\beta : \text{driving} \leftarrow \]
\[ F_3^\beta : \neg \text{alcohol} \leftarrow \]
A walk through the example

The decision theory

\[
\begin{align*}
R_{32}^\delta & : g_3 \leftarrow \text{drug, driving, } D(a_2) \\
R_{32}'^\delta & : g_3 \leftarrow \text{alcohol, driving, } D(a_2) \\
R_{41}^\delta & : g_4 \leftarrow D(a_1) \\
R_{51}^\delta & : g_5 \leftarrow D(a_1) \\
R_{61}^\delta & : g_6 \leftarrow D(a_1) \\
R_{62}^\delta & : g_6 \leftarrow D(a_2) \\
R_{72}^\delta & : g_7 \leftarrow D(a_2) \\
R_{31}^\delta & : g_3 \leftarrow D(a_1) \\
R_{31}'^\delta & : g_3 \leftarrow D(a_1) \\
R_{42}^\delta & : g_4 \leftarrow D(a_2) \\
R_{52}^\delta & : g_5 \leftarrow D(a_2) \\
R_{71}^\delta & : g_7 \leftarrow D(a_1)
\end{align*}
\]
Recursive and abductive argument

**Definition ([Vreeswijk 97, Dung, Kowalski & Toni 06])**

An argument \( A = \langle \text{conc}, \text{premise}, \text{asm} \rangle \) is:

1. **hypothetical**, i.e. built upon an assumption
   \[ \text{sent}(A) = \text{asm}(A) \]
   \[ \text{eg } A = \langle D(a_1), \emptyset, [D(a_1)] \rangle \text{ or } A = \langle \text{drug}, \emptyset, [\text{drug}] \rangle \]

2. **trivial**, i.e. built upon an unconditional ground statement
   \[ \text{sent}(A) = \text{premise}(A) \]
   \[ \text{eg } A = \langle \text{alcohol}, [\text{alcohol}], \emptyset \rangle \text{ or } A = \langle \neg \text{alcohol}, [\neg \text{alcohol}], \emptyset \rangle \]

3. a minimal and consistent **tree**, i.e. built upon a top rule where all literals in the body are the conclusions of subargument s.a:
   - \[ \text{sent}(A) = \bigcup_{A_i = \text{subarg}(A)} \text{sent}(A_i) \cup \text{conc}(A) \]
   - \[ \text{conc}(A) \not\subseteq \bigcup_{A_i = \text{subarg}(A)} \text{sent}(A_i) \text{ and } \neg \mathcal{I} (\text{sent}(A)) \]
   \[ \text{eg } A = \langle g_0, [g_1, g_2], [D(a_2), \text{drug}] \rangle \text{ or } A = \langle g_2, [g_6], [D(a_1)] \rangle \]
Interactions between concurrent or conflicting explanations

Definition (Attack relation)

attacks \((A, B)\) iff \(\text{sent}(A) \vdash \text{sent}(B)\).

Definition (Strength relation)

\(\succ^A\) is a (partial or total) preorder on arguments s.a.:

1. hypothetical arguments \(\succ^A\) trivial arguments \(\succ^A\) tree arguments
2. if \(\text{top}(A) \prec \text{top}(B)\), then \(A \succ^A B\);

Definition (Defeat relation)

\(A\) defeats \(B\)

1. attacks \((A, B)\)
2. \(\neg (B \succ^A A)\).
Determining whether a solution is ultimately suggested

Definition ([Dung 95])

1. \( A \) is **acceptable** wrt \( S \) iff \( \forall B \in A \) that defeats \( A \), \( S \) defeats \( B \).
2. \( S \) is **conflict-free** iff no argument in \( S \) is defeated by an argument in \( S \).
3. \( S \) is **admissible** iff \( S \) is conflict-free and every argument in \( S \) is acceptable wrt \( S \).
4. \( S \) is a **preferred extension** if it is a maximal admissible subset of \( A \).
5. An argument is **justified** if it is in every preferred extension.
6. An argument is **defensible** if it is in some but not all preferred extension.
Summary

Concrete argumentation framework for multi-attribute decision making

- Draw the influence diagram for decision analysis.
  *Hedgehogs know one big thing.*

- Recursive arguments are built on it.
  *Foxes knows many little things.*

- Preferences/Credibility/Uncertainty over the data structures.
  *Hedgehogs have intuitions.*

- Arguments attacks, have different strengths, defeat.
  *Foxes interact.*

- Deeper arguments with fewer assumptions have priority.
  *Hedgehogs never surrender.*

- Suggestion is determinated by an argumentation process.
  *Foxes change their minds.*
References

Philip E. Tetlock
Expert Political Judgment: How Good is It? How Can We Know?

R. T. Clemen.
Making Hard Decisions.

D. H. Berman and C. D. Hafner.
The potential of artificial intelligence to help solve the crisis in our legal system.
References (cont.)

P. M. Dung.
On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. 

H. Prakken and G. Sartor.
Argument-based logic programming with defeasible priorities. 

G. Vreeswijk.
Abstract argumentation systems. 

Phan Minh Dung, Robert A. Kowalski, Francesca Toni
Dialectic proof procedures for assumption-based, admissible argumentation 