Five weaknesses of ASPIC+

Leila Amgoud

amgoud@irit.fr

- Argumentation = an activity of reason aimed to increase (or decrease) the acceptability of a controversial standpoint by putting forward arguments
- Argumentation in AI = used for
 - reasoning about inconsistent premises
 - making decisions
 - modeling dialogues
 - ...

ASPIC+ (Prakken 2010) = an argumentation system
 It instantiates Dung's abstract framework

 Aim = { to show five serious flaws of ASPIC+ to study the properties of its underlying logics

- Argumentation = an activity of reason aimed to increase (or decrease) the acceptability of a controversial standpoint by putting forward arguments
- Argumentation in AI = used for
 - reasoning about inconsistent premises
 - making decisions
 - modeling dialogues
 - ...

• ASPIC+ (Prakken 2010) = an argumentation system

• It instantiates Dung's abstract framework

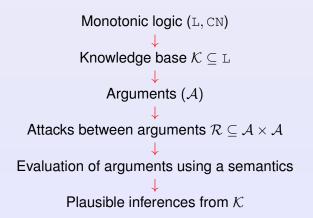
• Aim = to show five serious flaws of ASPIC+ to study the properties of its underlying logics

- Argumentation = an activity of reason aimed to increase (or decrease) the acceptability of a controversial standpoint by putting forward arguments
- Argumentation in AI = used for
 - reasoning about inconsistent premises
 - making decisions
 - modeling dialogues
 - ...
- ASPIC+ (Prakken 2010) = an argumentation system
 - It instantiates Dung's abstract framework

• Aim = to show five serious flaws of ASPIC+ to study the properties of its underlying logics

- Argumentation = an activity of reason aimed to increase (or decrease) the acceptability of a controversial standpoint by putting forward arguments
- Argumentation in AI = used for
 - reasoning about inconsistent premises
 - making decisions
 - modeling dialogues
 - ...
- ASPIC+ (Prakken 2010) = an argumentation system
 - It instantiates Dung's abstract framework

• Aim = to show five serious flaws of ASPIC+ to study the properties of its underlying logics



ASPIC+: Logical language

• Abstract logical language \mathcal{L} (for knowledge and names of rules)

• Strict / Defeasible rules: let $x_1, \ldots, x_n, x \in \mathcal{L}$

- *x*₁,..., *x*_n → *x* (if *x*₁,..., *x*_n hold then without exception *x* holds) *x*₁,..., *x*_n ⇒ *x* (if *x*₁,..., *x*_n hold then presumably *x* holds)
- They may represent either knowledge or reasoning patterns
- Contrariness function: $\overline{}: \mathcal{L} \mapsto 2^{\mathcal{L}}$. Let $x \in \overline{y}$.
 - if $y \notin \bar{x}$, then x is a contrary of y
 - otherwise, x and y are contradictory

Consistency: A set X ⊆ L is consistent iff ∄ x, y ∈ X s.t. x ∈ ȳ. Otherwise, X is inconsistent.

- Abstract logical language \mathcal{L} (for knowledge and names of rules)
- Strict / Defeasible rules: let $x_1, \ldots, x_n, x \in \mathcal{L}$
 - *x*₁,..., *x*_n → *x* (if *x*₁,..., *x*_n hold then without exception *x* holds)
 *x*₁,..., *x*_n ⇒ *x* (if *x*₁,..., *x*_n hold then presumably *x* holds)

 - They may represent either knowledge or reasoning patterns
- Contrariness function: $\overline{}: \mathcal{L} \longmapsto 2^{\mathcal{L}}$. Let $x \in \overline{y}$.
 - if $y \notin \overline{x}$, then x is a contrary of y
 - otherwise, x and y are contradictory

• Consistency: A set $X \subseteq \mathcal{L}$ is consistent iff $\nexists x, y \in X$ s.t. $x \in \overline{y}$. Otherwise, X is inconsistent.

- Abstract logical language \mathcal{L} (for knowledge and names of rules)
- Strict / Defeasible rules: let $x_1, \ldots, x_n, x \in \mathcal{L}$
 - $x_1, \ldots, x_n \to x$ (if x_1, \ldots, x_n hold then without exception x holds)
 - $x_1, \ldots, x_n \Rightarrow x$ (if x_1, \ldots, x_n hold then presumably x holds)
 - They may represent either knowledge or reasoning patterns
- Contrariness function: $\overline{}: \mathcal{L} \longmapsto 2^{\mathcal{L}}$. Let $x \in \overline{y}$.
 - if $y \notin \overline{x}$, then x is a contrary of y
 - otherwise, x and y are contradictory

• Consistency: A set $X \subseteq \mathcal{L}$ is consistent iff $\nexists x, y \in X$ s.t. $x \in \overline{y}$. Otherwise, X is inconsistent.

- Abstract logical language \mathcal{L} (for knowledge and names of rules)
- Strict / Defeasible rules: let $x_1, \ldots, x_n, x \in \mathcal{L}$
 - $x_1, \ldots, x_n \to x$ (if x_1, \ldots, x_n hold then without exception x holds)
 - $x_1, \ldots, x_n \Rightarrow x$ (if x_1, \ldots, x_n hold then presumably x holds)
 - They may represent either knowledge or reasoning patterns
- Contrariness function: $\bar{}: \mathcal{L} \longmapsto 2^{\mathcal{L}}$. Let $x \in \bar{y}$.
 - if $y \notin \bar{x}$, then x is a contrary of y
 - otherwise, x and y are contradictory
- Consistency: A set $X \subseteq \mathcal{L}$ is consistent iff $\nexists x, y \in X$ s.t. $x \in \overline{y}$. Otherwise, X is inconsistent.

- Abstract logical language L (for knowledge and names of rules)
- Strict / Defeasible rules: let $x_1, \ldots, x_n, x \in \mathcal{L}$
 - $x_1, \ldots, x_n \to x$ (if x_1, \ldots, x_n hold then without exception x holds)
 - $x_1, \ldots, x_n \Rightarrow x$ (if x_1, \ldots, x_n hold then presumably x holds)
 - They may represent either knowledge or reasoning patterns
- Contrariness function: $\bar{}: \mathcal{L} \longmapsto 2^{\mathcal{L}}$. Let $x \in \bar{y}$.
 - if $y \notin \bar{x}$, then x is a contrary of y
 - otherwise, x and y are contradictory
- Consistency: A set $X \subseteq \mathcal{L}$ is consistent iff $\nexists x, y \in X$ s.t. $x \in \overline{y}$. Otherwise, X is inconsistent.

Some remarks on the logical formalism (1/2)

• No restrictions on \mathcal{L} and rules. Thus,

- $x \rightarrow (y \rightarrow z)$ is a strict rule
- $(a \rightarrow b) \Rightarrow (x \rightarrow y)$ is a defeasible rule

No distinction between knowledge and names of defeasible rules ¬*f* ∈ *L* may be the name of *b* ⇒ *f* (birds generally fly)

The logical formalism is flawed

Some remarks on the logical formalism (1/2)

• No restrictions on \mathcal{L} and rules. Thus,

- $x \rightarrow (y \rightarrow z)$ is a strict rule
- $(a \rightarrow b) \Rightarrow (x \rightarrow y)$ is a defeasible rule

• No distinction between knowledge and names of defeasible rules

• $\neg f \in \mathcal{L}$ may be the name of $b \Rightarrow f$ (birds generally fly)

Conclusion

The logical formalism is flawed.

Some remarks on the logical formalism (1/2)

• No restrictions on $\mathcal L$ and rules. Thus,

- $x \rightarrow (y \rightarrow z)$ is a strict rule
- $(a \rightarrow b) \Rightarrow (x \rightarrow y)$ is a defeasible rule
- No distinction between knowledge and names of defeasible rules
 - $\neg f \in \mathcal{L}$ may be the name of $b \Rightarrow f$ (birds generally fly)

Conclusion

The logical formalism is flawed.

Some remarks on the logical formalism (2/2)

- Let \mathcal{L} be a propositional language
- Let ⁻ stand for classical negation
- \mathcal{R}_s = the inference patterns of propositional logic, $\mathcal{R}_d = \emptyset$
- The set $X = \{x, x \rightarrow y, \neg y\}$ is consistent in ASPIC+

Conclusion

- The semantics of the logical formalism is ambiguous.
- The logical formalism cannot capture classical logics.

Some remarks on the logical formalism (2/2)

- Let \mathcal{L} be a propositional language
- Let ⁻ stand for classical negation
- \mathcal{R}_s = the inference patterns of propositional logic, $\mathcal{R}_d = \emptyset$
- The set $X = \{x, x \rightarrow y, \neg y\}$ is consistent in ASPIC+

Conclusion

- The semantics of the logical formalism is ambiguous.
- The logical formalism cannot capture classical logics.

• Four bases: $\mathcal{K} = \mathcal{K}_n \cup \mathcal{K}_p \cup \mathcal{K}_a \cup \mathcal{K}_i$ s.t.

- \mathcal{K}_n : a set of axioms
- *K_p*: a set of ordinary premises
- \mathcal{K}_a : a set of assumptions
- *K_i*: a set of issues

• Remark: Strict and defeasible rules encode knowledge

- "Penguins do not fly" is a strict rule $(p \rightarrow \neg f)$ or an axiom?
- "Birds fly" is a defeasible rule $(b \Rightarrow f)$ or an ordinary premise?

• Four bases: $\mathcal{K} = \mathcal{K}_n \cup \mathcal{K}_p \cup \mathcal{K}_a \cup \mathcal{K}_i$ s.t.

- \mathcal{K}_n : a set of axioms
- *K_p*: a set of ordinary premises
- \mathcal{K}_a : a set of assumptions
- *K_i*: a set of issues

Remark: Strict and defeasible rules encode knowledge

- "Penguins do not fly" is a strict rule $(p \rightarrow \neg f)$ or an axiom?
- "Birds fly" is a defeasible rule $(b \Rightarrow f)$ or an ordinary premise?

Arguments

Arguments are trees

• Examples:

- *L*: a propositional language
- $K_p = \{x, y\}$ and $\mathcal{K}_n = \mathcal{K}_a = \mathcal{K}_i = \emptyset$
- $\mathcal{R}_s = \{x \to z\}$ and $\mathcal{R}_d = \{y, z \Rightarrow t\}$

• $x, x \rightarrow z$ is an argument in favor of z

• $x, x \rightarrow z, y, yz \Rightarrow t$ is an argument in favor of t

Conclusion

 ASPIC+ may miss intuitive conclusions Example:

- Let L be a propositional language and rules encode knowledge
- $\mathcal{K}_{p} = \{x \land y\}$ and $\mathcal{R}_{s} = \{x
 ightarrow z\}$
- No argument in favor of z. Thus, z will not be inferred!!

Arguments

Arguments are trees

• Examples:

- L: a propositional language
- $K_p = \{x, y\}$ and $\mathcal{K}_n = \mathcal{K}_a = \mathcal{K}_i = \emptyset$
- $\mathcal{R}_s = \{x \to z\}$ and $\mathcal{R}_d = \{y, z \Rightarrow t\}$
- $x, x \rightarrow z$ is an argument in favor of z
- $x, x \rightarrow z, y, yz \Rightarrow t$ is an argument in favor of t

Conclusion

ASPIC+ may miss intuitive conclusions

Example:

- Let L be a propositional language and rules encode knowledge
- $\mathcal{K}_p = \{x \land y\}$ and $\mathcal{R}_s = \{x \to z\}$
- No argument in favor of z. Thus, z will not be inferred!!

Arguments

Arguments are trees

• Examples:

- L: a propositional language
- $K_p = \{x, y\}$ and $\mathcal{K}_n = \mathcal{K}_a = \mathcal{K}_i = \emptyset$
- $\mathcal{R}_s = \{x \to z\}$ and $\mathcal{R}_d = \{y, z \Rightarrow t\}$
- $x, x \rightarrow z$ is an argument in favor of z
- $x, x \rightarrow z, y, yz \Rightarrow t$ is an argument in favor of t

Conclusion

ASPIC+ may miss intuitive conclusions

Example:

- $\bullet~$ Let ${\mathcal L}$ be a propositional language and rules encode knowledge
- $\mathcal{K}_{p} = \{x \land y\}$ and $\mathcal{R}_{s} = \{x \rightarrow z\}$
- No argument in favor of z. Thus, z will not be inferred!!

• Rebutting: to undermine the conclusion of an argument

• $A: t, t \Rightarrow z, z \Rightarrow x, x \rightarrow y$ rebuts $B: t', t' \rightarrow z', z' \rightarrow x', x' \Rightarrow \neg y$

- B does not rebut A
- But, A is not more certain than B!

Conclusion ASPIC+ builds on counter-intuitive assumptions • Rebutting: to undermine the conclusion of an argument

• $A: t, t \Rightarrow z, z \Rightarrow x, x \rightarrow y$ rebuts $B: t', t' \rightarrow z', z' \rightarrow x', x' \Rightarrow \neg y$

• B does not rebut A

• But, A is not more certain than B!

Conclusion

ASPIC+ builds on counter-intuitive assumptions.

• Rebutting: to undermine the conclusion of an argument

• $A: t, t \Rightarrow z, z \Rightarrow x, x \rightarrow y$ rebuts $B: t', t' \rightarrow z', z' \rightarrow x', x' \Rightarrow \neg y$

• B does not rebut A

• But, A is not more certain than B!

Conclusion

ASPIC+ builds on counter-intuitive assumptions.

Undermining: to undermine a premise of an argument x, x → z undermines ¬z, ¬z → v

Undercutting: to undermine the applicability of a defeasible rule
Let K_n = {b, ¬f}, R_d = {b ⇒ f} where ¬f is the name of b ⇒ f
A: b
B: b, b ⇒ f
C: ¬f

B undercuts itself, B undermines C and C rebuts B
 The system infers b and ¬f!

Conclusion

• $x, x \rightarrow z$ undermines $\neg z, \neg z \rightarrow v$

• Undercutting: to undermine the applicability of a defeasible rule

• Let $\mathcal{K}_n = \{b, \neg f\}, \mathcal{R}_d = \{b \Rightarrow f\}$ where $\neg f$ is the name of $b \Rightarrow f$

A: bB: $b, b \Rightarrow f$ C: $\neg f$

B undercuts itself, B undermines C and C rebuts B
 The system infers b and ¬f!

Conclusion

• $x, x \rightarrow z$ undermines $\neg z, \neg z \rightarrow v$

Undercutting: to undermine the applicability of a defeasible rule
Let K_n = {b, ¬f}, R_d = {b ⇒ f} where ¬f is the name of b ⇒ f
A: b
B: b, b ⇒ f
C: ¬f

B undercuts itself, *B* undermines *C* and *C* rebuts *B*The system infers *b* and ¬*f*!

Conclusion

- $x, x \rightarrow z$ undermines $\neg z, \neg z \rightarrow v$
- Undercutting: to undermine the applicability of a defeasible rule
 Let K_n = {b, ¬f}, R_d = {b ⇒ f} where ¬f is the name of b ⇒ f
 A: b
 B: b, b ⇒ f
 C: ¬f
 - *B* undercuts itself, *B* undermines *C* and *C* rebuts *B*The system infers *b* and ¬*f*!

Conclusion

- $x, x \rightarrow z$ undermines $\neg z, \neg z \rightarrow v$
- Undercutting: to undermine the applicability of a defeasible rule
 Let K_n = {b, ¬f}, R_d = {b ⇒ f} where ¬f is the name of b ⇒ f
 A: b
 B: b, b ⇒ f
 C: ¬f
 - B undercuts itself, B undermines C and C rebuts B
 - The system infers b and $\neg f!$

Conclusion

- $x, x \rightarrow z$ undermines $\neg z, \neg z \rightarrow v$
- Undercutting: to undermine the applicability of a defeasible rule
 Let K_n = {b, ¬f}, R_d = {b ⇒ f} where ¬f is the name of b ⇒ f
 A: b
 B: b, b ⇒ f
 C: ¬f
 - B undercuts itself, B undermines C and C rebuts B
 - The system infers b and $\neg f!$

Conclusion

- Dung's acceptability semantics (Dung, 1995)
 - E.g. Preferred semantics: maximal non-conflicting and self-defending sets of arguments

• Let
$$\mathcal{R}_d = \{ \Rightarrow a, \Rightarrow b, \Rightarrow x, \Rightarrow z, a \Rightarrow (x \rightarrow y), b \Rightarrow (z \rightarrow \neg y) \},\ \mathcal{R}_s = \mathcal{K}_n = \mathcal{K}_p = \mathcal{K}_a = \mathcal{K}_i = \emptyset$$

$$\begin{array}{ccc} \text{D:} \Rightarrow z & \text{E:} \Rightarrow z, a \Rightarrow (x \rightarrow y) & \text{F:} \Rightarrow b, b \Rightarrow (z \rightarrow \neg y) \end{array}$$

- $\{A, B, C, D, E, F\}$ is the unique preferred extension
- $a, b, x, z, x \rightarrow y, z \rightarrow \neg y$ are outputs of the system
- The output is not closed (y is not inferred)
- The output is indirectly inconsistent (y and $\neg y$)

Conclusion

- Dung's acceptability semantics (Dung, 1995)
 - E.g. Preferred semantics: maximal non-conflicting and self-defending sets of arguments
- Let $\mathcal{R}_d = \{\Rightarrow a, \Rightarrow b, \Rightarrow x, \Rightarrow z, a \Rightarrow (x \rightarrow y), b \Rightarrow (z \rightarrow \neg y)\},\ \mathcal{R}_s = \mathcal{K}_n = \mathcal{K}_p = \mathcal{K}_a = \mathcal{K}_i = \emptyset$
 - $\begin{array}{lll} A: \Rightarrow a & B: \Rightarrow b & C: \Rightarrow x \\ D: \Rightarrow z & E: \Rightarrow a, a \Rightarrow (x \rightarrow v) & F: \Rightarrow b, b \Rightarrow (z \rightarrow v) \end{array}$
 - $\{A, B, C, D, E, F\}$ is the unique preferred extension
 - $a, b, x, z, x \rightarrow y, z \rightarrow \neg y$ are outputs of the system
 - The output is not closed (y is not inferred)
 - The output is indirectly inconsistent (y and $\neg y$)

Conclusion

- Dung's acceptability semantics (Dung, 1995)
 - E.g. Preferred semantics: maximal non-conflicting and self-defending sets of arguments

• Let
$$\mathcal{R}_d = \{ \Rightarrow a, \Rightarrow b, \Rightarrow x, \Rightarrow z, a \Rightarrow (x \rightarrow y), b \Rightarrow (z \rightarrow \neg y) \},\$$

 $\mathcal{R}_s = \mathcal{K}_n = \mathcal{K}_p = \mathcal{K}_a = \mathcal{K}_i = \emptyset$
A: $\Rightarrow a$ B: $\Rightarrow b$ C: $\Rightarrow x$
D: $\Rightarrow z$ E: $\Rightarrow a, a \Rightarrow (x \rightarrow y)$ F: $\Rightarrow b, b \Rightarrow (z \rightarrow \neg y)$

• $\{A, B, C, D, E, F\}$ is the unique preferred extension

- $a, b, x, z, x \rightarrow y, z \rightarrow \neg y$ are outputs of the system
- The output is not closed (y is not inferred)
- The output is indirectly inconsistent (y and $\neg y$)

Conclusion

- Dung's acceptability semantics (Dung, 1995)
 - E.g. Preferred semantics: maximal non-conflicting and self-defending sets of arguments

• Let
$$\mathcal{R}_d = \{ \Rightarrow a, \Rightarrow b, \Rightarrow x, \Rightarrow z, a \Rightarrow (x \rightarrow y), b \Rightarrow (z \rightarrow \neg y) \},\ \mathcal{R}_s = \mathcal{K}_n = \mathcal{K}_p = \mathcal{K}_a = \mathcal{K}_i = \emptyset$$

$$A: \Rightarrow a \qquad B: \Rightarrow b \qquad C: \Rightarrow x$$

 $\mathsf{D} :\Rightarrow z \qquad \mathsf{E} :\Rightarrow a, a \Rightarrow (x \rightarrow y) \qquad \mathsf{F} :\Rightarrow b, b \Rightarrow (z \rightarrow \neg y)$

• {*A*, *B*, *C*, *D*, *E*, *F*} is the unique preferred extension

• $a, b, x, z, x \rightarrow y, z \rightarrow \neg y$ are outputs of the system

- The output is not closed (*y* is not inferred)
- The output is indirectly inconsistent (y and $\neg y$)

Conclusion

- Dung's acceptability semantics (Dung, 1995)
 - E.g. Preferred semantics: maximal non-conflicting and self-defending sets of arguments

• Let
$$\mathcal{R}_d = \{\Rightarrow a, \Rightarrow b, \Rightarrow x, \Rightarrow z, a \Rightarrow (x \rightarrow y), b \Rightarrow (z \rightarrow \neg y)\},\ \mathcal{R}_s = \mathcal{K}_n = \mathcal{K}_p = \mathcal{K}_a = \mathcal{K}_i = \emptyset$$

A: $\Rightarrow a$ B: $\Rightarrow b$ C: $\Rightarrow x$

$$\mathsf{D}:\Rightarrow z \qquad \mathsf{E}:\Rightarrow a,a\Rightarrow (x\rightarrow y) \qquad \mathsf{F}:\Rightarrow b,b\Rightarrow (z\rightarrow \neg y)$$

• {*A*, *B*, *C*, *D*, *E*, *F*} is the unique preferred extension

- $a, b, x, z, x \rightarrow y, z \rightarrow \neg y$ are outputs of the system
- The output is not closed (*y* is not inferred)
- The output is indirectly inconsistent (y and $\neg y$)

Conclusion

- Dung's acceptability semantics (Dung, 1995)
 - E.g. Preferred semantics: maximal non-conflicting and self-defending sets of arguments

• Let
$$\mathcal{R}_d = \{ \Rightarrow a, \Rightarrow b, \Rightarrow x, \Rightarrow z, a \Rightarrow (x \rightarrow y), b \Rightarrow (z \rightarrow \neg y) \},\ \mathcal{R}_s = \mathcal{K}_n = \mathcal{K}_p = \mathcal{K}_a = \mathcal{K}_i = \emptyset$$

$$\begin{array}{lll} \mathsf{A}:\Rightarrow a & \mathsf{B}:\Rightarrow b & \mathsf{C}:\Rightarrow x \\ \mathsf{D}:\Rightarrow z & \mathsf{E}:\Rightarrow a, a\Rightarrow (x\rightarrow y) & \mathsf{F}:\Rightarrow b, b\Rightarrow (z\rightarrow \neg y) \end{array}$$

- {*A*, *B*, *C*, *D*, *E*, *F*} is the unique preferred extension
- $a, b, x, z, x \rightarrow y, z \rightarrow \neg y$ are outputs of the system
- The output is not closed (y is not inferred)
- The output is indirectly inconsistent (y and $\neg y$)

Conclusion

- Dung's acceptability semantics (Dung, 1995)
 - E.g. Preferred semantics: maximal non-conflicting and self-defending sets of arguments

• Let
$$\mathcal{R}_d = \{ \Rightarrow a, \Rightarrow b, \Rightarrow x, \Rightarrow z, a \Rightarrow (x \rightarrow y), b \Rightarrow (z \rightarrow \neg y) \},\ \mathcal{R}_s = \mathcal{K}_n = \mathcal{K}_p = \mathcal{K}_a = \mathcal{K}_i = \emptyset$$

$$\begin{array}{lll} \mathsf{A}\!:\,\Rightarrow\,a & \mathsf{B}\!:\,\Rightarrow\,b & \mathsf{C}\!:\,\Rightarrow\,x \\ \mathsf{D}\!:\,\Rightarrow\,z & \mathsf{E}\!:\,\Rightarrow\,a,a\,\Rightarrow\,(x\rightarrow y) & \mathsf{F}\!:\,\Rightarrow\,b,b\,\Rightarrow\,(z\rightarrow\neg y) \end{array}$$

- $\{A, B, C, D, E, F\}$ is the unique preferred extension
- $a, b, x, z, x \rightarrow y, z \rightarrow \neg y$ are outputs of the system
- The output is not closed (y is not inferred)
- The output is indirectly inconsistent (y and $\neg y$)

Conclusion

ASPIC+ violates the basic rationality postulates.

Amgoud (IRIT)

ASPIC+ suffers from five main problems:

- its logical formalism is ill-defined
- it may return undesirable results
- it builds on some counter-intuitive assumptions
- it violates some rationality postulates
- it allows counter-intuitive instantiations