Logic for AI

Master 1 IFI

Andrea G. B. Tettamanzi

Nice Sophia Antipolis University

Computer Science Department

andrea.tettamanzi@unice.fr

Session 8

Argumentation

Agenda

- Introduction
- Argumentation Frameworks
- Semantics

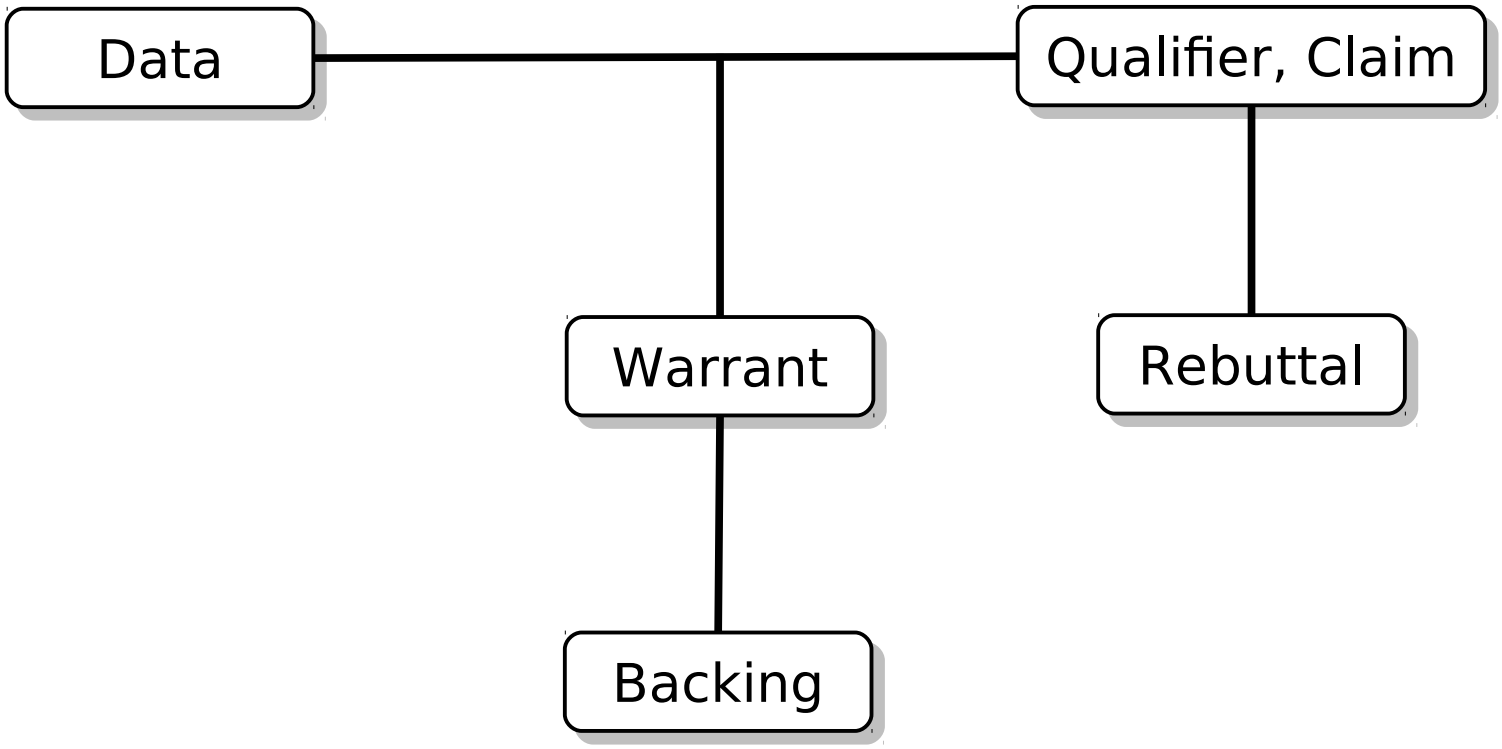
Introduction

- Argumentation is the interdisciplinary study of how conclusions can be reached through logical reasoning
- In AI: a tool to provide a proof-theoretic semantics for non-monotonic logic
- Non-monotonic logic is any formal framework devised to capture and represent *defeasible* inference
  - The reasoner draws conclusions tentatively, reserving the right to retract them in the light of further information
  - Dealing with conflicts
  - We touched upon this notion in the lecture on Belief Revision
  - Another noteworthy formalism is default logic, which we will not cover

Why do we argue?

- **Information-seeking:** an agent seeks to answer some question(s) with the help of another agent, who knows the answer
- **Inquiry:** agents collaborate to answer a question, whose answer they do not know
- **Persuasion:** an agent seeks to persuade another to accept a proposition they do not currently endorse
- **Negotiation:** bargaining over allocation of resources
- **Deliberation:** decide which action(s) should be adopted in a given situation
- **Eristic:** verbal quarrel rather than physical fighting to solve a dispute

Toulmin Model of Arguments



Example of an Argument

Congress should ban animal research (Claim #1) because animals are tortured in experiments that have no necessary benefit for humans such as the testing of cosmetics (Data).

The well being of animals is more important than the profits of the cosmetics industry (Warrant). Only congress has the authority to make such a law (Warrant) because the corporations can simply move from state to state to avoid legal penalties (Backing).

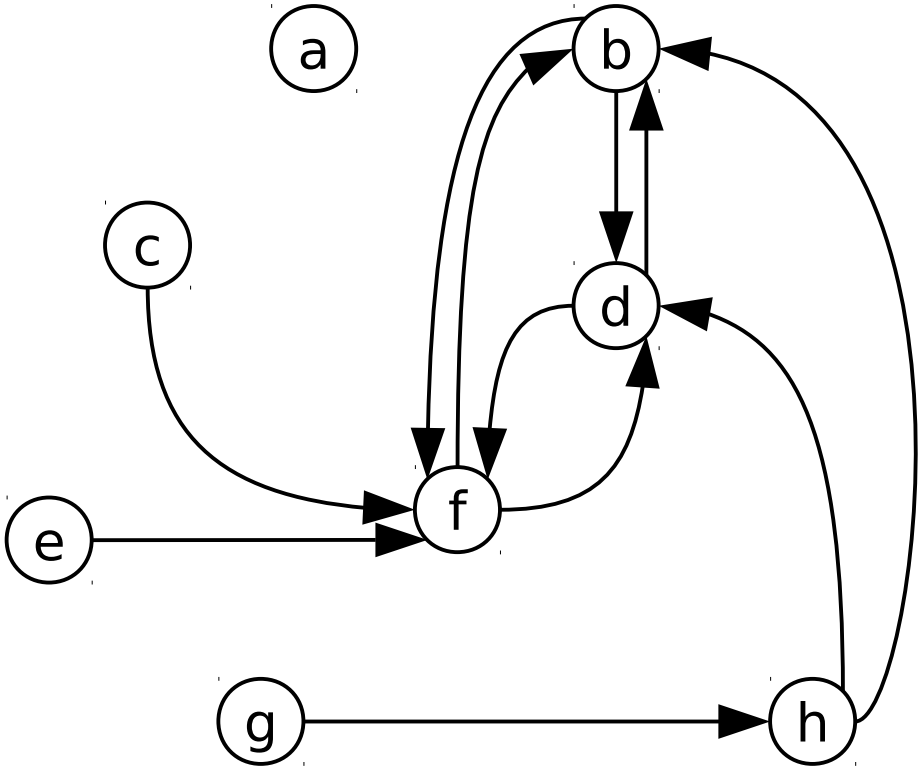
Of course, this ban should not apply to medical research (Qualifier). A law to ban all research would go too far (Rebuttal).

So, the law would probably (Qualifier) have to be carefully written to define the kinds of research intended (Claim #2).

Abstract Argumentation

- Proposed by Phan Minh Dung at IJCAI 1993
- Basic ideas:
  - Disregard the internal structure of arguments
  - Consider only how they attack each other
- An argumentation framework is defined as a pair <*A*, *attacks*>
  - *A* is a set of arguments (abstract elements)
  - *attacks* is a binary relation on *A*, the attack relation
- An argumentation framework can be viewed as a graph

Example



Semantics

- A **semantics** for an argumentation framework is a way to identify sets of arguments “surviving the conflict together”
- What this intuitive notion means exactly depends on the particular semantics
- Semantics of argumentation frameworks can be stated as
  - Extensions (sets of accepted arguments)
  - Labelings (mappings assigning labels to arguments)

Semantics Properties

- A set *S* of arguments is **conflict-free** if there are no arguments *a* and *b* in *S* such that *a attacks b*
- An argument a in A is **acceptable** w.r.t. a set S of arguments iff for each argument b in A, if *b attacks a*, then b is attacked by S
- A conflict-free set of arguments S is **admissible** iff each argument in S is acceptable w.r.t. S
- **strongly admissible** iff every argument defended by S is in S
- An extension S is **i-maximal** iff no proper subset of S is an extension

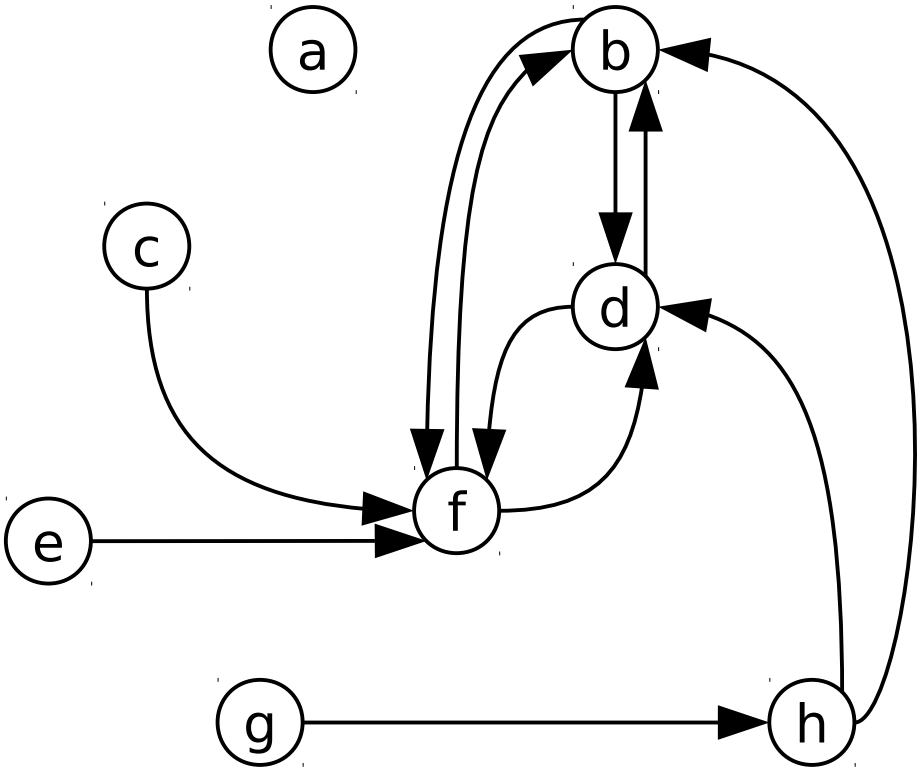
Complete Extension

- Admissible (thus conflict-free)
- Each defended argument is included (reinstatement)

- Intuitively, the notion of complete extensions captures the kind of confident rational agent who believes in everything it can defend.

Complete Extensions

$$\{\ \{a, c, d, e, g\},\ \{a, b, c, e, g\},\ \{a, c, e, g\}\ \}$$



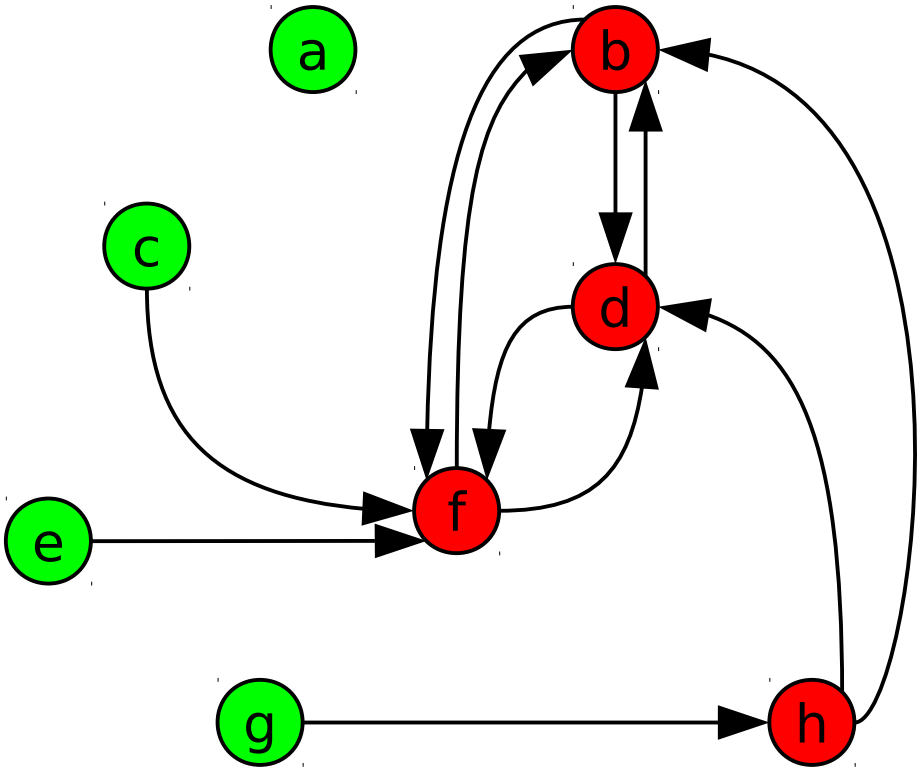
Grounded Extension

- Strongly admissible (thus conflict-free and admissible)
- Minimum complete extension

- Grounded extensions are “skeptical”

Grounded Extensions

$\{ \{a, c, e, g\} \}$



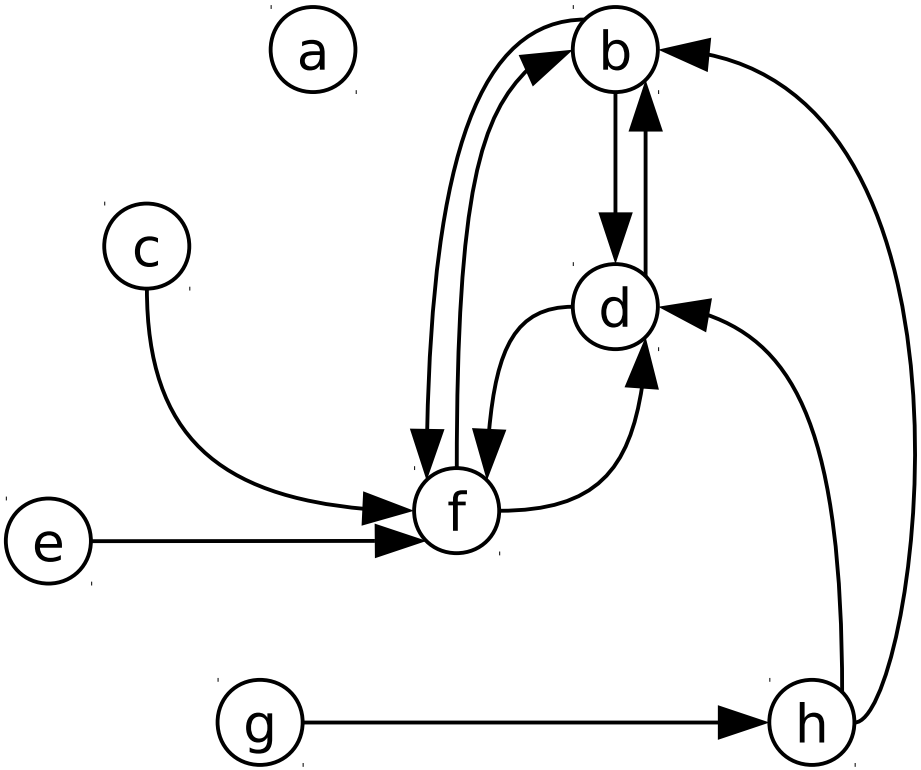
Preferred Extension

- Admissible (thus conflict-free)
- Maximal (w.r.t. set inclusion)
- Maximum complete extensions

- Preferred extensions are “credulous”

Preferred Extensions

$$\{\ \{a, c, d, e, g\},\ \{a, b, c, e, g\}\ \}$$



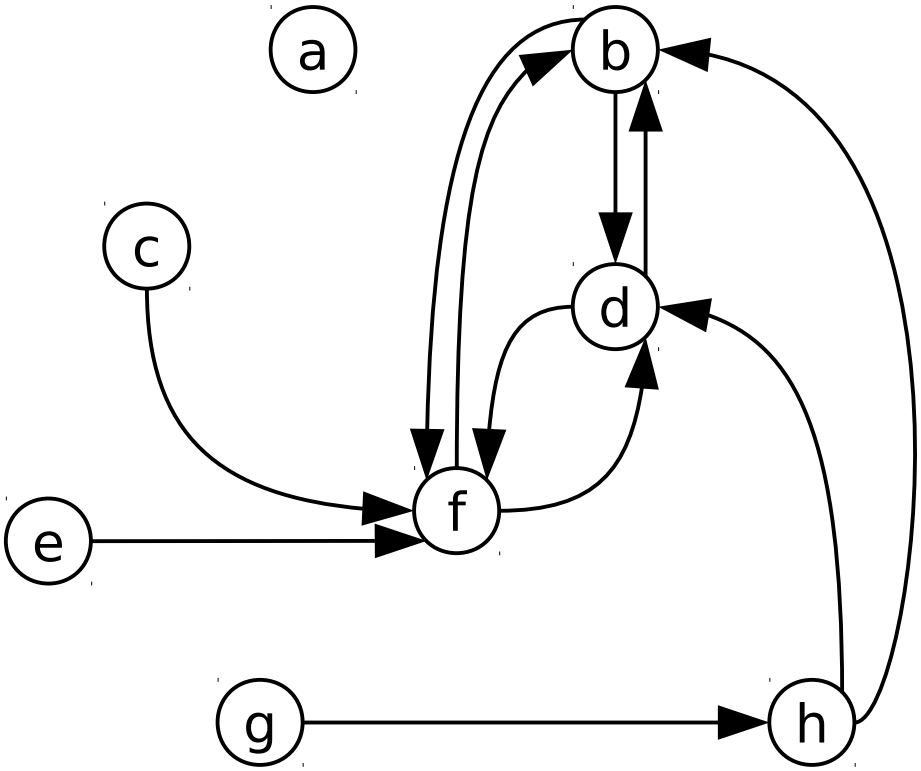
Stable Extension

- Complete extension
- Attacking all the arguments outside
- The absence of odd-length cycles is a sufficient condition for the existence of stable extensions

- Every stable extension is a preferred extension
  - but not *vice versa*

Stable Extensions

$$\{\ \{a, c, d, e, g\},\ \{a, b, c, e, g\}\ \}$$



Coherence

- An argumentation framework *AF* is **coherent** if each preferred extension of *AF* is also stable
- An argumentation framework *AF* is **relatively grounded** if its grounded extension coincides with the intersection of all preferred extensions
- There exists at least one stable extension in a coherent argumentation framework

Complete Labelings

- Arguments are labeled as { IN, OUT, UNDEC }
  - An argument is IN if all of its attackers are OUT
  - An argument is OUT if at least one of its attackers is IN
  - An argument is UNDEC otherwise

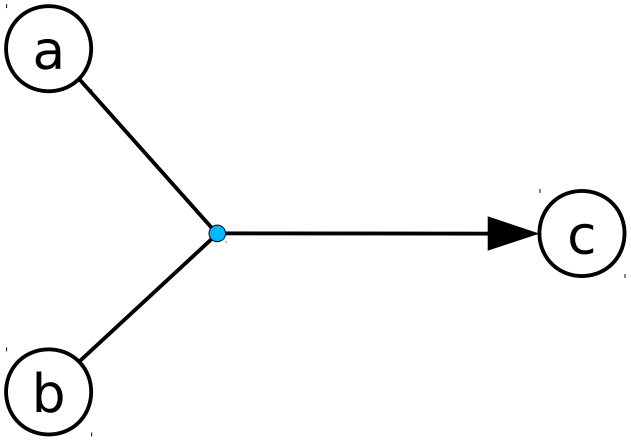
- Maximize UNDEC = Grounded Extension
- Maximize IN = Preferred Extension
- No UNDEC = Stable Extension

Extending Dung's Framework

- Dung's framework captures negative interactions between arguments
- However, it does not capture several intuitive properties of human argumentation
  - Joint attack
  - Recursive/meta-arguments
  - Preferences
  - Support
  - Argument strength

Joint Attack

- Both *a* and *b* must be accepted in order for *c* not to be accepted
- All the previous results and definitions map directly
- Only the definition of *attacks* needs modification



Preference-Based Argumentation

- Witness *a* claims P, witness *b* claims ¬P, but *a* is more reliable than *b*
- A preference-based argumentation framework (PAF) is a triple <A, attacks, ≥ >, where ≥ is a partial ordering over A
- "*a* ≥ *b*" states that a is preferred to b
- An AF is transformed into a PAF by shifting from the notion of attack to that of defeat
- *a defeats b* iff *a attacks b* and *a* ≥ *b*

Strength

- Humans often claim that some arguments are stronger than others
- Such strength can come from
  - its internal structure—the validity of the inference pattern to check the tenability of the claim
  - its social consensus (e.g., the number of favorable and unfavorable votes)
  - the authority of the source (or the “reasoner”) offering it
    - May be a measure of the reliability of the source, like competence, expertise, trust, reputation, and the like

Fuzzy Labeling

- Let <*A*, $\rightarrow$> be an abstract argumentation framework
- A fuzzy labeling is a total function $\alpha : A \rightarrow [0, 1]$

$$\alpha(a) \leq 1 - \max_{b:b \rightarrow a} \alpha(b)$$

- In addition arguments may have a “strength” in [0, 1]
  - Trustworthiness of their source, support, etc.
- In that case, *A* may be viewed as a fuzzy set
- A fuzzy reinstatement labeling is a fuzzy labeling such that

$$\alpha(a) = \min\{A(a), 1 - \max_{b:b \rightarrow a} \alpha(b)\}$$

Computing a Fuzzy Reinstatement Labeling

- We define the sequence

$$\alpha_0(a) = A(a)$$

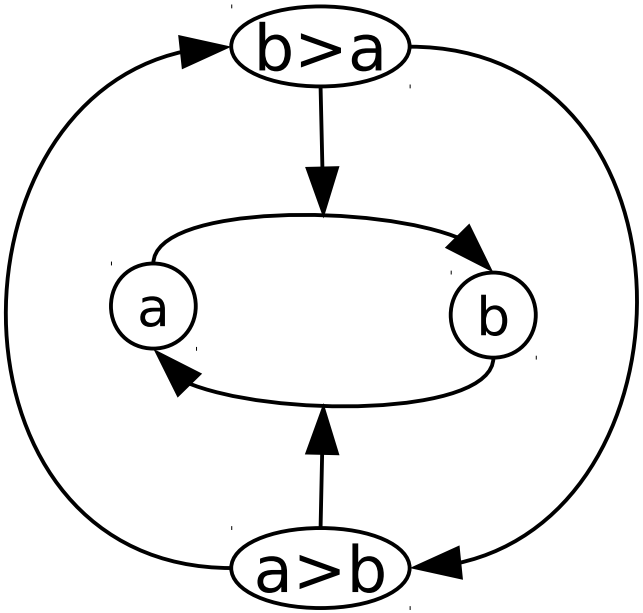
$$\alpha_{t+1}(a) = \frac{1}{2}\alpha_t(a) + \frac{1}{2}\min\{A(a), 1 - \max_{b:b\to a} \alpha_t(b)\}$$

- This sequence always converges
- Its limit is a fuzzy reinstatement labeling

$$\alpha(a) = \lim_{t\to\infty} \alpha_t(a)$$

Extended Frameworks

- The idea of these frameworks is to allow attacks on attacks
- Capturing preferences, undercuts, and the like in a natural way



Bipolar Argumentation

- Attacks between arguments allow for reinstatements to occur, allowing arguments to defend one another
- Arguments can also build on top of one another, or strengthen each other through support
- Bipolar argumentation frameworks allow arguments to interact by either attacking or supporting one another

< *A*, *attacks*, *supports* >

- Different formalisms treat support differently

Thank you for your attention