Multiagent Systems

14. Argumentation

B. Nebel, C. Becker-Asano, S. Wölfl

Albert-Ludwigs-Universität Freiburg

July 23, 2014

Where are we?

- Bargaining
- Alternating offers
- Negotiation decision functions
- Task-oriented domains
- Bargaining for resource allocation

Today ...

- Argumentation in Multiagent Systems

14.1 Motivation

Argumentation

- Agents may have mutually **contradicting beliefs**:
  I believe p; you believe $\neg p$
  I believe p; from p follows q; you believe $\neg q$
- How can agents reach agreements about **what to believe**?
- **Argumentation** provides principled techniques for deciding what to believe in the face of inconsistencies
- We achieve this by comparing arguments that can be compiled from the agents' beliefs
- Arguments usually present beliefs and describe reasonable justifications

What is an argument?

Intuitively, an argument consists of:

- a **claim**
- a set of reasons for the claim (**justification**, **support**)

Different types of arguments:

- **Rebutting argument**: an argument that claims the negation of another argument
- **Undercutting argument**: an argument with a claim that contradicts some assumption used to justify another argument
- **Counterargument**: Given some argument, a counterargument rebuts or undercuts the argument

Modes of arguments

At least four different modes of arguments can be identified between humans (Gilbert, 1994):

- **Logical mode**: deductive, proof-like, concerned with making correct inferences
- **Emotional mode**: appeals to feelings, attitudes, etc.
- **Visceral mode**: physical, social aspects
- **Kisceral mode**: appeals to the intuitive, mystical or religious

$\rightsquigarrow$ Different types are used/accepted in different situations
(e.g. no emotional or kisceral mode arguments allowed in courts of law)

14.2 Abstract Argumentation

Abstract argumentation system

We can decide what to believe while looking at arguments at the abstract level (Dung, 1995):

- Disregarding internal structures of arguments
- Focus on the attack relation between arguments $(a, b, c, d, \ldots)$: a **attacks** b or $a \rightarrow b$
- Not concerned with the origin of arguments or the attack relation

Abstract argumentation system

An **abstract argumentation system** $A = \langle X, \rightarrow \rangle$ is defined by:

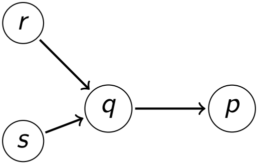
- a set of **arguments** X,
- a binary attack relation on arguments $\rightarrow \subseteq X \times X$.

Example

Consider the following argumentation system:

$$\langle \{p, q, r, s\}, \{(r, q), (s, q), (q, p)\} \rangle,$$

i.e., with arguments: p, q, r, s, and attacks: $r \rightarrow q$, $s \rightarrow q$, $q \rightarrow p$.



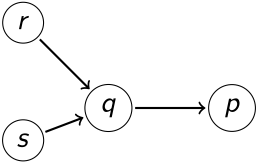
$\rightsquigarrow$ Which sets of arguments can be considered **rationally justified**?

Conditions for argument sets

Consider a Dung-style argumentation system (as in the definition).

- A set of arguments S is **conflict-free** if there is no pair of arguments $a, b \in S$ such that $a \to b$.
- An argument a is **acceptable** with respect to a set S of arguments if each argument a' that attacks a is attacked by some argument in S.
- A conflict-free set of arguments S is **admissible** if each argument in S is acceptable wrt. S.

Example (cont'd)



- The following argument sets are conflict-free:

$$\emptyset, \{p\}, \{q\}, \{r\}, \{s\}, \{r,s\}, \{p,r\}, \{p,s\}, \{p,r,s\}.$$

- The following argument sets are admissible:

$$\emptyset, \{r\}, \{s\}, \{r,s\}, \{p,r\}, \{p,s\}, \{p,r,s\}.$$

Preferred extensions

Given a Dung-style argumentation system.

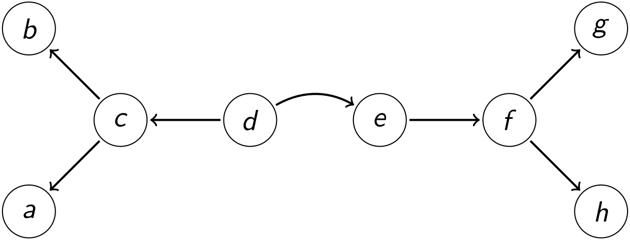
- An admissible set of arguments is called **preferred extension** if it is maximal (wrt. set inclusion).
- An argument is **sceptically accepted** if it is contained in each preferred extension.
- An argument is **credulously accepted** if it is contained in some preferred extension.

Preferred extensions help determine which arguments should be accepted but are not always useful:

- . . . are not necessarily unique,
- the only preferred extension may be the empty set

Nevertheless, each argumentation system has at least some preferred extension (note, preferred extension need not be non-empty).

Example



Which argument sets are preferred extensions?

Reasoning tasks in argumentation systems

Theorem

- *The problem to check whether a given set of arguments is admissible can be decided in polynomial time.*
- *The problem to check whether a given set of arguments is a preferred extension is coNP-complete.*
- *The problem to check whether a given argument is contained in some preferred extension is NP-complete.*
- *The problem to check whether a given argumentation system has a stable extension is NP-complete (a* **stable** *extension is a set of arguments S such that each argument not in S is attacked by some argument in S).*

Grounded extensions

An alternative notion of acceptability: the notion of **grounded extension**.

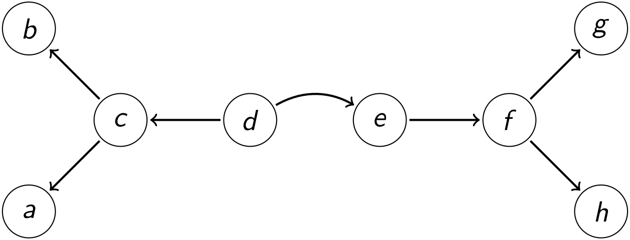
Grounded extension

Given an abstract argumentation system $\mathcal{A} = \langle V, \rightarrow \rangle$, the **grounded extension** in $\mathcal{A}$ is incrementally built as follows:

1. Mark all arguments that are not attacked as "in".
2. Mark all arguments as "out" which are attacked by some argument marked as "in".
3. Set $V := V \setminus \{\text{"out"-nodes"}\}, \rightarrow := \rightarrow \cap V \times V$.
4. Iterate until the argumentation graph does not change.

- The grounded extension always exists and is guaranteed to be unique, but
- . . . may be empty (if no argument is not attacked initially)

Example



Compute the grounded extension?

Grounded extensions (fix-point characterization)

Let $\mathcal{A} = \langle X, \rightarrow \rangle$ be an abstract argumentation system with finite X. Consider the following function:

$$F\colon 2^X \rightarrow 2^X, S \mapsto \{a \in X\colon a \text{ is acceptable wrt. } S\}$$

- The grounded extension of an argumentation system is the least fix-point of the function F.
- Consider the sequence:

$$\begin{aligned} E_0 &:= \emptyset \\ E_{i+1} &:= \{a \in X\colon a \text{ is acceptable wrt. } S\} \end{aligned}$$

Then $E = \bigcup E_i$ is the grounded extension of $\mathcal{A}$.

Limitations of abstract argumentation systems

- In abstract argumentation systems all arguments are equally strong — which is not very realistic
  $\rightsquigarrow$ **Preference-based argumentation systems** (e.g., Amgoud et al. 1998f) model preference (weights) of arguments.
- Acceptability of arguments can depend on the target audience (e.g., newspaper vs. scientific article)
  $\rightsquigarrow$ **Value-based argumentation systems** (Bench-Capon et. al, 2003ff)
- Arguments in abstract argumentation systems do not have an internal (logical) structure
  $\rightsquigarrow$ **Deductive argumentation systems**

14.3 Deductive Argumentation Systems

Deductive Argumentation Systems

The "purest", most rational kind of argument: in classical logic, argument = sequence of inferences leading to a conclusion
Write $\Gamma \vdash \varphi$ to denote that some sequence of inference steps from premises in Γ will allow us to establish proposition φ

Deductive argument

Let K be a set of formulae (intuitively, the formulae accepted by all participants of an argumentation, not necessarily consistent).
A **deductive argument** is a pair (Γ, ϕ) where:

- $\Gamma \subseteq K$
- $\Gamma \vdash \varphi$
- Γ is logically consistent
- Γ is minimal (i.e. no proper subset of Γ satisfies these conditions)

Argument types

Some important types of arguments:

- **Tautological arguments**: (Γ, φ) with $\Gamma = \emptyset$
- **Non-trivial arguments**: (Γ, φ) with $\Gamma \neq \emptyset$
- **Rebutting argument**: (Γ, φ) rebuts (Γ', φ') if $\varphi \equiv \neg\varphi'$
- **Undercutting argument**: (Γ, φ) undercuts (Γ', φ') if $\varphi \equiv \neg\gamma$ for some $\gamma \in \Gamma'$
- **Defeating argument**: (Γ, φ) defeats against (Γ', φ') if it rebuts or undercuts the latter.

Example

Consider the following example:

$$
\begin{aligned}
\mathrm{Arg}_1 &:= \big(\{\mathrm{human}(\mathrm{Heracles}), \mathrm{human}(X) \rightarrow \mathrm{mortal}(X)\}, \\
&\quad \mathrm{mortal}(\mathrm{Heracles})\big) \\
\mathrm{Arg}_2 &:= \big(\{\mathrm{father}(\mathrm{Heracles}, \mathrm{Zeus}), \mathrm{father}(X, \mathrm{Zeus}) \rightarrow \mathrm{divine}(X), \\
&\quad\quad \mathrm{divine}(X) \rightarrow \neg\mathrm{mortal}(X)\}, \\
&\quad \neg\mathrm{mortal}(\mathrm{Heracles})\big) \\
\mathrm{Arg}_3 &:= \big(\{\neg(\mathrm{father}(X, \mathrm{Zeus}) \rightarrow \mathrm{divine}(X))\}, \\
&\quad \neg(\mathrm{father}(X, \mathrm{Zeus}) \rightarrow \mathrm{divine}(X))\big)
\end{aligned}
$$

- Arg_1 and Arg_2 are mutually rebutting
- Arg_3 undercuts Arg_2

Which arguments are stronger, more acceptable?

Argument Classes

We can identify five classes of argument type in order of increasing acceptability:

- A1: The class of all arguments that can be constructed
- A2: The class of all **non-trivial** arguments that can be constructed
- A3: The class of all arguments that can be constructed with **no rebutting arguments**
- A4: The class of all arguments that can be constructed with **no undercutting arguments**
- A5: The class of all **tautological arguments** that can be constructed

Example: Argument classes

- Arguments Arg_1 and Arg_2 are in (A2) (mutually rebutting)
- Argument

$$(\emptyset, \text{divine}(\text{Heracles}) \vee \neg\text{divine}(\text{Heracles}))$$

  is in (A5).
- Argument

$$\big(\{\text{father}(\text{Apollo}, \text{Zeus}), \text{father}(X, \text{Zeus}) \rightarrow \text{divine}(X), \text{divine}(X) \rightarrow \neg\text{mortal}(X)\}, \neg\text{mortal}(\text{Apollo})\big)$$

  is in (A4).

14.4 Argumentation-based Dialogue Systems

Argumentation dialogue systems

Agents engage in dialogue to convince other agents of some state of affairs. Consider two agents 0 and 1 engaging in the following dialogue:

- Agent 0 attempts to convince 1 of some argument
- Agent 1 attempts to rebut or undercut it
- Agent 0 in turn attempts to defeat 1's argument
- and so on ...

Each steps in such a dialogue is a **move** $(\text{Player}, \text{Arg})$ (with $\text{Player} \in \{0, 1\}, \text{Arg} \in A(DB)$)

A **dialogue history** is a sequence of moves $(m_0, \ldots, m_k)$ s.t.:

- $\text{Player}_{2i} = 0$, $\text{Player}_{2i+1} = 1$ for all $i \geq 0$
- If $\text{Player}_i = \text{Player}_j$ and $i \neq j$, then $\text{Arg}_i \neq \text{Arg}_j$
- Arg_{i+1} defeats Arg_i for all $i \geq 0$

A dialogue **ends** if no further moves are possible, the **winner** then is Player_k.

Types of dialogue

Typology due to Walton and Krabbe (1995):

Type	Initial situation	Main goal	Participants' aim
Persuasion	conflict of opinion	resolve the issue	persuade other
Negotiation	conflict of interest	make a deal	get best deal
Inquiry	general ignorance	growth of knowledge	find a proof
Deliberation	need for action	reach a decision	influence outcome
Information seeking	personal ignorance	spread knowledge	gain or pass on knowledge
Eristics	conflict/ antagonism	reaching an accommodation	strike other party
Mixed	various	various	various

14.5 Summary

- Thanks

Summary

- Argumentation
- Abstract argumentation systems
- Deductive argumentation systems
- Argumentation-based dialogue
- **Next time:** Logics for Multiagent Systems

Acknowledgments

These lecture slides are based on the following resources:

- Dr. Michael Rovatsos, The University of Edinburgh
  http://www.inf.ed.ac.uk/teaching/courses/abs/abs-timetable.html
- Michael Wooldridge: **An Introduction to MultiAgent Systems**, John Wiley & Sons, 2nd edition 2009.
- Paul E. Dunne & T.J.M. Bench-Capon: Coherence in finite argument systems. In: **Artificial Intelligence** 141 (2002), p. 187–203.
- P. Besnard & A. Hunter, **Elements of Argumentation**, MIT Press, 2008.
- Simon Parsons, Carles Sierra, & Nick Jennings: Agents that reason and negotiate by arguing, In: **Journal of Logic and computation**, 8(3), pp. 261-292, 1998.