

# Online debating platforms: from argument mapping to decision-aiding

---

N. Maudet

June 2017



Joint work with many colleagues, within the AMANDE project  
TeD Summer School



## **Online debate platforms**

---

# Types of platform

Several types of platforms exist:

- **time-centric** contributions organized along a timeline (twitter, web-forums, etc.)
- **question-centric** open questions, contributions typically promoted by votes (stackoverflow, etc.)
- **topic-centric** contributions aggregated into a single coherent article, debates may occur in dedicated pages (wikis, etc.)

M. Klein. *A Critical Review of Crowd-Scale Online Deliberation Technologies*. MIT report.

## Types of platform

- **debate-centric** focus on controversial topics and binary decisions, list of PRO and CON arguments (debatepedia)
- **argument-centric** tree-structure of arguments, may split into several issues and sub-arguments (DebateGraph, Arguman, etc.)

M. Klein. *A Critical Review of Crowd-Scale Online Deliberation Technologies*. MIT report.

## What do you mean by arguments?

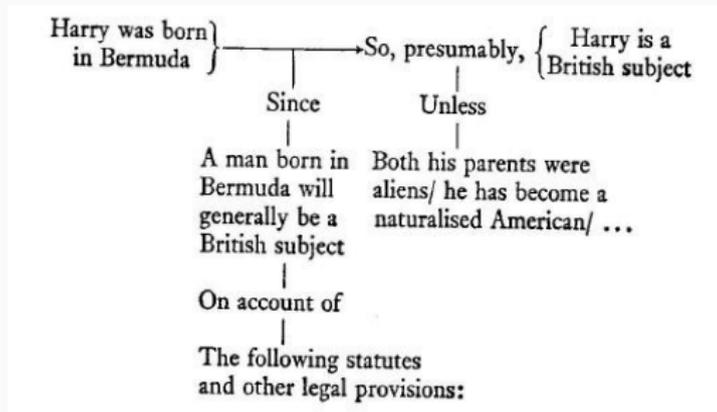
Argumentation has been studied for ages by linguists, philosophers, logicians...

Eg. Toulmin (argument scheme):

# What do you mean by arguments?

Argumentation has been studied for ages by linguists, philosophers, logicians...

Eg. Toulmin (argument scheme):



Many types of **argument schemes**: analogical, statistical, etc.

## What do you mean by arguments?

In online debate platforms the quality of arguments is, well, variable

# What do you mean by arguments?

In online debate platforms the quality of arguments is, well, variable

 **KingOfPopForever** (6902) 3 points  

Pluto is a planet and also a dog I like that's Gofy's.

6 years ago Side: Hell yeah! Pluto rocks!

[Support](#) [Dispute](#) [Clarify](#)

 **Merlin13** (1268) 2 points  

Well, it's referred to as a dwarf planet. Doesn't that still make it a planet of some sort since it's in its name? Or should they just refer to it as a small resemblance of a planet?

6 years ago Side: Hell yeah! Pluto rocks!

[Support](#) [Dispute](#) [Clarify](#)

[↑ Hide Replies](#)

 **Hellno** (17613) 4 points  

Well, I guess we have found several others in the same orbit that are the same size or ever bigger... But why did it take this long? I'm suspicious!

6 years ago Side: Hell yeah! Pluto rocks!

[Support](#) [Dispute](#) [Clarify](#)

[↑ Hide Replies](#)

 **Merlin13** (1268) 3 points  

It's all a conspiracy! LOL

 **Yani12** (12) 2 points  

Pluto was demoted from planetary status due to a number of prominent factors. The first was that, like Ceres it was surrounded by many objects of similar size and composition. Ceres was first thought to be a planet because of its spherical shape but later on it was labelled a dwarf planet due to its size, and the surrounding bodies were called asteroids.

Secondly Pluto's orbit is substantially different from those of the other planets. It is inclined and varies from being around 49AU to 30AU from the sun. It is also very small having a radius of 1180km, smaller then our moon. Charon, Pluto's moon, has a radius of around 600km. Which initiated another debate over whether or not it should be called a binary planetary system. Astronomers soon discovered many other objects around Pluto such as Makemake and Sedna which were similar in size to that of Pluto. Then in early 2005 Eris was discovered and it was larger than Pluto. This led to the revaluation of the definition of 'planet' in 2006 by the IAU. The traditional definition of a planet (a celestial body moving around a star) was changed, and for an object to be classified as a planet it had to;

- Be in orbit around the Sun,
- Have sufficient mass to assume hydrostatic equilibrium
- Have had "cleared the neighbourhood" around its orbit.

If an object was found but only met the first two points it would be classified as a dwarf planet. Pluto, Eris and the other asteroids are now considered part of the Kuiper belt.

5 years ago Side: No! Pluto is a dump!

[Support](#) [Dispute](#) [Clarify](#)

(Taken from: CreateDebate, debate "Is Pluto a Planet Or Not?")

# Are all arguments acceptable in a debate?

Some forms of argumentation have been identified as non-acceptable in the course of a debate: **fallacies**.

The most common fallacies have been identified and listed.

▶ fallacies

See in particular the work of D. Walton.

D. Walton. *Fallacies: Selected Papers: 1972-1982*. Studies in Logic, 2007.

# Overview of the talk

Online debate platforms

Examples of debate platforms

Analysis of debates

Decision-aiding

## **Examples of debate platforms**

---

## Example: Debatepedia

- a controversial **issue** is posted
- the issue may be split into several **sub-issues**
- arguments are listed side by side: PRO vs. CON
- fair amount of debates but the site seems inactive for a while

## Example: Argüman

- a controversial **issue** is posted
- tree of arguments, linked with 3 relations **but**, **however**, and **because**.
- users can **vote** on arguments they support
- users can tag arguments as **fallacies**

## Example: DebateGraph

- a controversial **issue** is posted
- issues can be splitted into **sub-issues**
- positions can be taken on the different issues
- supporting arguments, attacking arguments, and groups of arguments can be created
- arguments can be given a **score** on a 1-10 scale

## Example: ChangeMyView

- an **original poster** (OP) puts forward an opinion on an issue
- other users submit contributions aiming at convincing the OP
- a **delta system** is used to acknowledge a change of opinion
- **moderators** enforce some rules
- users may **vote** on comments,

## Example: ChangeMyView

- an **original poster** (OP) puts forward an opinion on an issue
- other users submit contributions aiming at convincing the OP
- a **delta system** is used to acknowledge a change of opinion
- **moderators** enforce some rules
- users may **vote** on comments, **but**:
  - ☞ *CMW interestingly discourages the use of negative votes.*
  - ☞ *More generally, the guidelines state: "Vote based on the quality of the comment, not whether you agree or not."*

## Example: ChangeMyView

Some (interesting) specificities of CMW:

- a **large community** (221,000 subscribers as of 2015)
- many users respond to the OP, allowing **diversity** of style, arguments, etc.
- **explicit persuasion acknowledgment**, at the level of arguments.
- openness of the OP to change her mind, so potentially **many positive examples** of actual persuasion

## Summary: requested features and issues

Easy-to-use interface, with a **handful of simple relations**

**but**

- users may not be competent enough to distinguish them
- on the other hand the system may not allow to express relations that may seem natural

## Summary: requested features and issues

Easy-to-use interface, with a **handful of simple relations**

**but**

- users may not be competent enough to distinguish them
- on the other hand the system may not allow to express relations that may seem natural

Possibility to **vote on arguments**

**but**

- what is the semantics of a vote on an arguments?  
(I believe it is true, I support it, I would like to see it accepted)

# **Analysis of debates**

---

## Descriptive objectives

Starting from the data available on online debates, we may ask some questions, eg:

1. Are users good at specifying arguments and relations?
2. Do online debates exhibit specific structures?
3. What makes arguments persuasive?

I will mostly build on 3 papers, illustrating different approaches and techniques: expert annotations, graph analysis, NLP tools.

[CV] Cabrio and Villata. *Towards a Benchmark of Natural Language Arguments*. ArXiv1405.0941.

[CPRST] Cerutti et al. *A pilot study in using argumentation for online debates*. SAFA-16.

[TNDL] Tan et al. *Winning Arguments: Interaction Dynamics and Persuasion Strategies in Good-faith Online Discussions*. WWW-16.

**Disclaimer:** all figures are borrowed from these papers.

## Q1: Quality of users' annotated argumentative data

Evidence from the analysis of debates held on Debatepedia and ProCon reported in [CV].

### Methodology

- construction of (bipolar) argumentation graphs : gold standard obtained by expert annotators
- NLP and argumentation mining techniques (textual entailment) used to infer **attack** and **support** relations

## Q1: Quality of users' annotated argumentative data

Evidence from the analysis of debates held on Debatepedia and ProCon reported in [CV].

### Methodology

- construction of (bipolar) argumentation graphs : gold standard obtained by expert annotators
- NLP and argumentation mining techniques (textual entailment) used to infer **attack** and **support** relations

### Conclusions

*“abstract argumentation might not be fully appropriate since such natural language arguments have (possibly complex) internal structures and may include sub-arguments”*

# Q1: Quality of users' annotated argumentative data

Evidence from the analysis of a debate created on CreateDebate reported in [CPRST].

## Methodology

- on the platform users are asked to tag their posts as **support**, **dispute**, **clarification**
- ad-hoc annotation scheme consisting in
  - 5 types of nodes: question, answer, standard statement, partial statement, distractor statement; and
  - 6 types of edges: **answer**, **explicit/implicit support/attack**, **meta support/attack**, **node-to-edge support/attack**, **expansion**

## Q1: Quality of users' annotated argumentative data

Evidence from the analysis of a debate created on CreateDebate reported in [CPRST].

### Conclusions

- users only used the dispute relation (!)
- expert annotators made interesting use of non-standards relations (eg. expansion allowed to group arguments)

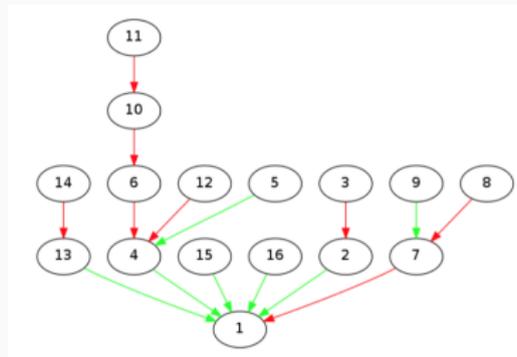
*“having non-expert annotations generated by debaters can be useful as a rough starting point for expert or automated annotation of the relation between arguments. Nevertheless, one needs to keep in mind that these non-expert annotations are biased and imperfect.”*

## Q2: What is the structure of debates?

The actual structure may differ from the one induced by the platform. Because of “mistakes” (see before) but also because there may be additional implicit relations.

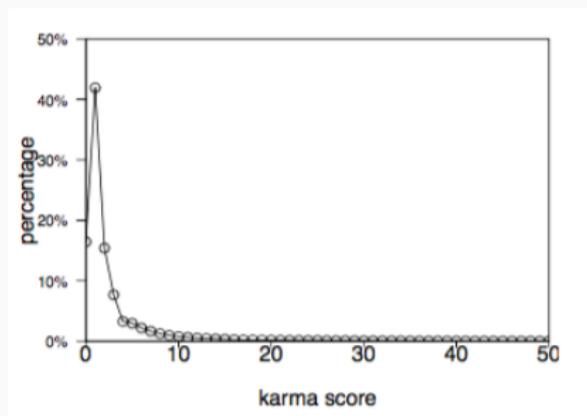
[CPRST] obtain a richer structure, involving in particular meta-level relations.

[CV] observe simple structures, without any occurrence of cycles



## Q2: What is the structure of debates?

Regarding **votes** and comments, unsurprisingly, a study performed on CMW emphasize that very few comments attract most upvotes



### Q3: Which are the most persuasive arguments?

This is a very difficult question, given the extreme scarcity of data [CPRST] note that:

*“We [...] lack a ground truth (for assessing which position debated is strongest)”*

In that respect, the study of [TNDL] on CMW is extremely interesting.

### Q3: Which are the most persuasive arguments?

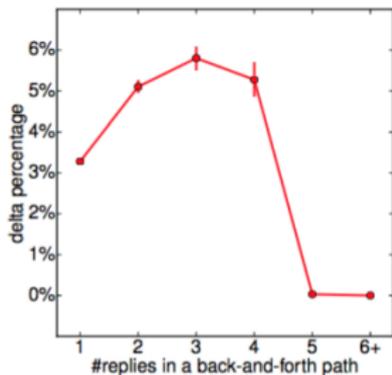
Evidence from CMW dataset: “how the interaction patterns in a debate relate to a challenger’s success?”

- **entry time** — relation between entry time of the challenger and chance of success: 🖱️ *decreases with entry time*

### Q3: Which are the most persuasive arguments?

Evidence from CMW dataset: “how the interaction patterns in a debate relate to a challenger’s success?”

- **entry time** — relation between entry time of the challenger and chance of success: 🖱️ *decreases with entry time*
- **back-and-forth** — sequences involving only the OP and a given challenger. 🖱️ *non-monotonic, and no chance of receiving a  $\Delta$  when length  $\geq 10$  (5 replies from the challenger)*



### Q3: Which are the most persuasive arguments?

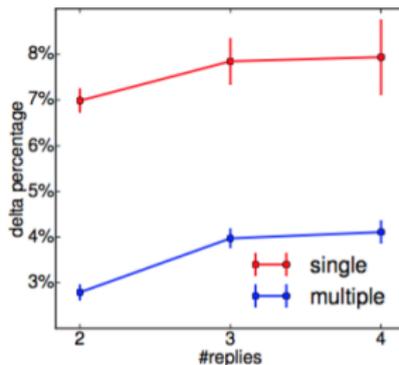
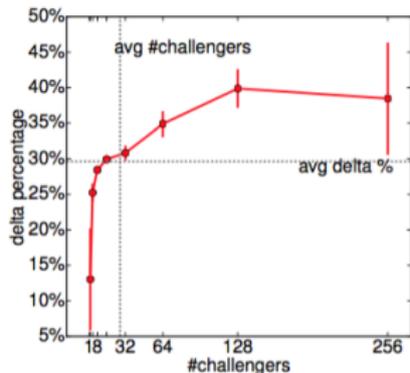
Evidence from CMW dataset: “how the interaction patterns in a debate relate to the likelihood of the OP to change her view”

- **number unique challengers** — grows with the number of unique challengers, up to a saturation level

### Q3: Which are the most persuasive arguments?

Evidence from CMW dataset: “how the interaction patterns in a debate relate to the likelihood of the OP to change her view”

- **number unique challengers** — grows with the number of unique challengers, up to a saturation level
- **single-challenger subtrees vs. multiple-challenger** — “when talking about the same counterargument, challengers might not be adding value to it, or they might even disagree”



## Decision-aiding

---

## Normative objectives

Suppose the decision lies in the hands of the moderator (not always the case). Here are some tasks for which the system may help:

1. deciding which side won (in binary issue debates)
2. extract “coherent” subparts of the debate (for summarizing, for setting up a team of representatives, etc.)
3. classify users in terms of their contributions
4. identify / diminish the influence of malicious users

## Q1: Which side won the debate?

We often we just want to know which of alternatives “yes” or “no” is the most supported.

In practice this is based on a **gradual evaluation** or **ranking semantics**, whose objective is to:

- assign **scores** to arguments, noted  $v(x)$ ; or
- simply **rank** (ordinally) the different arguments.

A basic but strong requirement that we have is that the evaluation (or ranking) is unique.

## Q1: Which side won the debate?

Suppose for a start the most basic model **argumentation systems** as a pair  $\langle Arg, \rightarrow \rangle$  where  $Arg$  is a set of arguments,  $\rightarrow$  is a binary relation representing attack relations among arguments

The **categorizer function** of Besnard and Hunter:

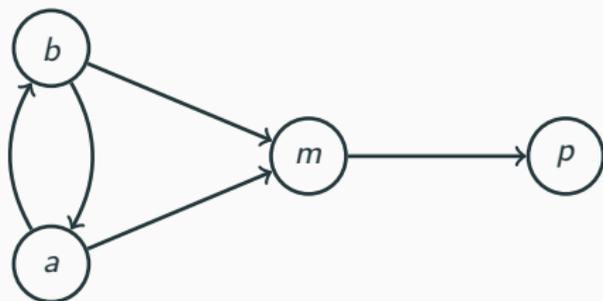
$$v(a) = \frac{1}{1 + \sum_{b \in Att(a)} v(b)}$$

Note: the **uniqueness** of solutions and **convergence** of algorithms is sometimes hard to guarantee for such fixed-point approaches.

Dung. *On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games.* AIJ-95.

Besnard, Hunter. *A logic-based theory of deductive arguments.* AIJ-2001.

## Q1: Which side won the debate?



Let us compute  $v(a)$  and  $v(b)$ . We have

$$v(a) = \frac{1}{1+v(b)} \text{ and } v(b) = \frac{1}{1+v(a)}.$$

By solving we get  $v(a) = v(b) = \frac{\sqrt{5}-1}{2} \simeq 0.68$ .

Then  $v(m) = \frac{1}{1+2 \times 0.68} \simeq 0.42$ . And finally  $v(p) \simeq 0.7$

## Q1: Which side won the debate?

But of course many other semantics may be conceived, including semantics accounting for **support only systems**, and for **bipolar argumentation structures** (attack and support)

Cayrol, Lagasquie. *Graduality in argumentation*. JAIR-05.

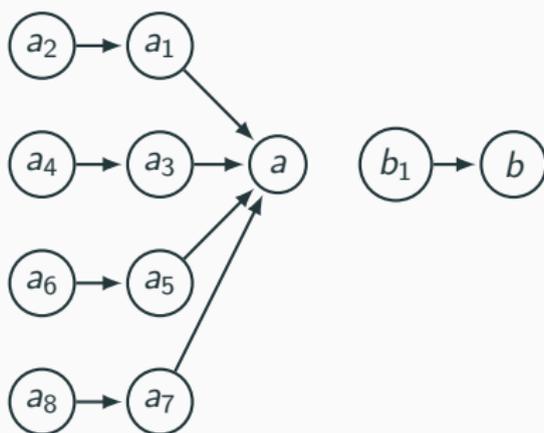
Cayrol, Lagasquie. *Gradual Valuation for Bipolar Argumentation Frameworks*. ESQUARU-05.

Amgoud *et al.* *Evaluation of arguments from support relations: Axioms and Semantics*. IJCAI-16.

## Q1: Which side won the debate?

So how do we decide which approach makes sense?

The **axiomatic approach** list properties that should be satisfied (or not) by ranking semantics, eg. *attack* vs. *full defense*



**Figure 1:** Should  $a$  be ranked higher than  $b$ ?

# Q1: Which side won the debate?

The axiomatic landscape of ranking semantics (cf. Hannu's talk):

Properties	SAF	Cat	Dbs	Bbs	$\alpha$ -Bbs	Tuples*	M&T	FL	IGD	CS	Propa <sub><math>\epsilon</math></sub>	Propa <sub>1+<math>\epsilon</math></sub>	Propa <sub>1-<math>\epsilon</math></sub>	Gr
Abs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
In	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓
VP	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×
DP	✓	✓	✓	✓	✓	×	×	×	×	✓	✓	✓	✓	×
CT	✓	✓	✓	✓	✓	×	×	(×)	(×)	✓	✓	✓	×	×
SCT	✓	✓	✓	✓	✓	×	×	×	×	✓	✓	✓	×	×
CP	×	×	✓	✓	×	×	×	×	×	×	×	×	×	×
QP	×	×	×	×	×	×	×	✓	×	×	×	×	×	✓
DDP	×	×	×	✓	×	×	×	✓	×	×	×	✓	✓	×
SC	×	×	×	×	×	-	✓	×	×	×	×	×	×	×
⊕DB	×	×	×	×	×	×	×	×	×	×	×	×	×	×
+DB	×	×	×	×	×	✓	×	×	×	×	×	×	✓	×
↑AB	✓	✓	✓	✓	(✓)	✓	×	×	×	(✓)	✓	✓	✓	×
↑DB	✓	✓	✓	✓	(✓)	✓	×	×	×	(✓)	✓	✓	✓	×
+AB	✓	✓	✓	✓	✓	✓	×	×	(✓)	✓	✓	✓	✓	×
Tot	✓	✓	✓	✓	✓	×	✓	✓	×	✓	✓	✓	✓	✓
NaE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AvsFD	×	×	×	×	×	✓	✓	✓	(×)	×	×	✓	✓	✓

Bonzon, Delobelle, Konieczny, Maudet. *A comparative study of ranking-based semantics for abstract argumentation*. AAI-16.

## Q1: Which side won the debate?

The relevance of these principles may depend on the application!

- in **persuasion pitches**, the phenomena of **protocatalepsis** says that it is more convincing to anticipate the (potential) counter-arguments of the audience.
- recall also the findings of CMW regarding back-and-forth sequences: showing ability to defend his position is good.

This violates the widely accepted principle of **void precedence**, saying that non-attacked arguments must have the higher ranks.

## Q1: Which side won the debate?

How to account for the votes of people on arguments:  $v^+$  for positive votes,  $v^-$  for negative votes.

First, we compute the ratio of positive votes expressed on  $a$ :

$$\tau(a) = \frac{v^+}{v^+ + v^- + \epsilon}$$

The **simple product semantics** accounts for this.

$$M(a) = \tau(a) \wedge \neg \vee \{M(a_i) : a_i \in \text{Att}(a)\}$$

avec  $\wedge$  product T-Norm et  $\vee$  probabilistic sum T-CoNorm.

Leite and Martins. *Social Argumentation Frameworks*. IJCAI-11.

## Q1: Which side won the debate?



We have  $v(a) = 4/5 \times (1 - v(b))$  and  $v(b) = 1/5 \times (1 - v(a))$ .  
By solving we get  $v(a) \simeq 0.76$  and  $v(b) \simeq 0.05$

## Q1: Which side won the debate?

The conjecture of uniqueness of models was eventually proven false for this semantics: from 4 arguments, there are argumentation systems exhibiting several solutions (but they involve cycles of course)

Other approaches recently proposed in the literature (like DF-QuAF and ArgDec, cf. Antonio's talk)

*Amgoud et al. A note on the uniqueness of models in social abstract argumentation. ArXiv.*

*Baroni et al. Automatic evaluation of design alternatives with quantitative argumentation. Argument and Computation, 2015.*

*Amgoud et al. Acceptability semantics for weighted argumentation frameworks. IJCAI-17.*

## Q2: Extracting coherent subparts of the debates

A crash course in Dung semantics.

The question is how to select **sets** of acceptable arguments. Each different method to perform this selection is an **argumentation semantics**.

- **conflict-freeness**— the set of arguments  $X$  must not contain internal contradiction, ie. there is no  $(x, y) \in X$  st.  $x \rightarrow y$ .

## Q2: Extracting coherent subparts of the debates

A crash course in Dung semantics.

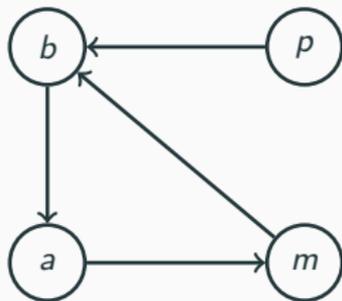
The question is how to select **sets** of acceptable arguments. Each different method to perform this selection is an **argumentation semantics**.

- **conflict-freeness**— the set of arguments  $X$  must not contain internal contradiction, ie. there is no  $(x, y) \in X$  st.  $x \rightarrow y$ .

From this principle we can already define:

- **naive semantics**— select  $\subseteq$ -maximal non-conflicting sets
- **stable semantics**— a set  $X$  is stable if any argument not in  $X$  is attacked by some argument in  $X$

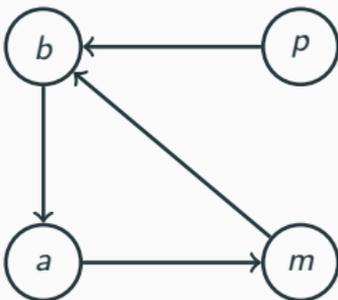
## Q2: Extracting coherent subparts of the debates



Naive:

Stable:

## Q2: Extracting coherent subparts of the debates



Naive:  $\{p, a\}, \{p, m\}$

Stable:  $\{p, a\}$

## Q2: Extracting coherent subparts of the debates

The **defense principle** captures the ability of a set of arguments to counter-attack any attacking argument.

- **defense**— a set  $S$  defends  $x$  if, for any  $y$  attacking  $x$ , there exists  $z \in S$  such that  $z$  attacks  $y$  (“ $z$  defends  $x$  against  $y$ ”)

## Q2: Extracting coherent subparts of the debates

The **defense principle** captures the ability of a set of arguments to counter-attack any attacking argument.

- **defense**— a set  $S$  defends  $x$  if, for any  $y$  attacking  $x$ , there exists  $z \in S$  such that  $z$  attacks  $y$  (“ $z$  defends  $x$  against  $y$ ”)

From this principle we can define :

- **admissible sets**— sets that defend their own arguments
- **admissible semantics**—  $\subseteq$ -maximal admissible sets

## Q2: Extracting coherent subparts of the debates

Maximality of those sets can be interpreted as a fixed-point:

- **defense function**— define  $F : S \rightarrow 2^{Arg}$  as the sets of arguments that are defended by the set  $S$

## Q2: Extracting coherent subparts of the debates

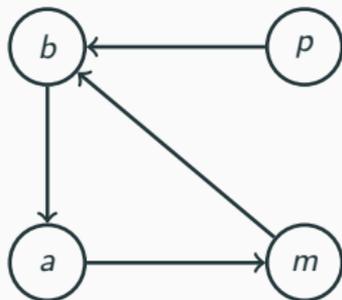
Maximality of those sets can be interpreted as a fixed-point:

- **defense function**— define  $F : S \rightarrow 2^{Arg}$  as the sets of arguments that are defended by the set  $S$

From this principle we can define :

- **complete semantics**— sets that include all the arguments they defend, ie. sets  $S$  that are fixed-points of the defense function, ie.  $F(S) = S$
- **grounded semantics**— least fixed point of the defense function, ie. apply iteratively  $F(\emptyset)$

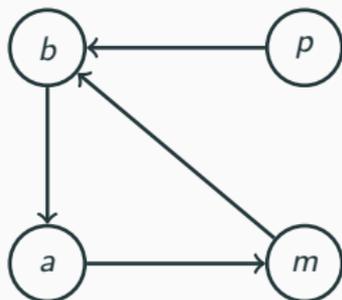
## Q2: Extracting coherent subparts of the debates



Admissible sets:

Grounded, preferred:

## Q2: Extracting coherent subparts of the debates

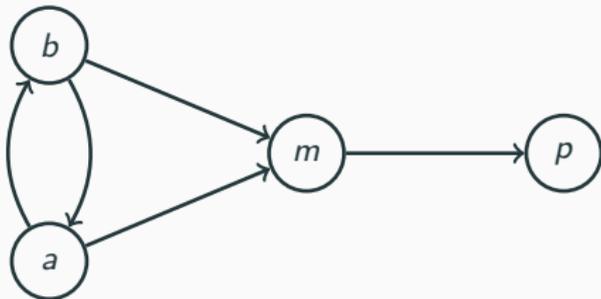


Admissible sets:  $\{\}, \{p\}, \{p, a\}$

Grounded, preferred:  $\{p, a\}$

## Q2: Extracting coherent subparts of the debates

The **floating conclusion** example :



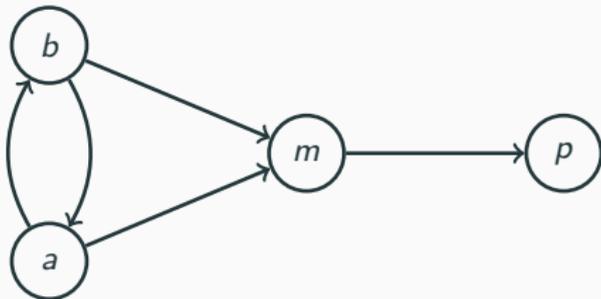
Admissible sets:

Preferred, stable:

Grounded:

## Q2: Extracting coherent subparts of the debates

The **floating conclusion** example :



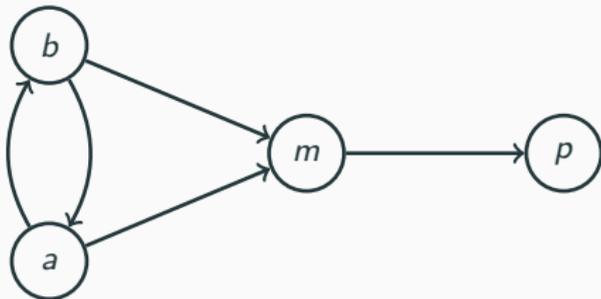
Admissible sets:  $\{\}, \{a\}, \{b\}, \{a, p\}, \{b, p\}$

Preferred, stable:

Grounded:

## Q2: Extracting coherent subparts of the debates

The **floating conclusion** example :



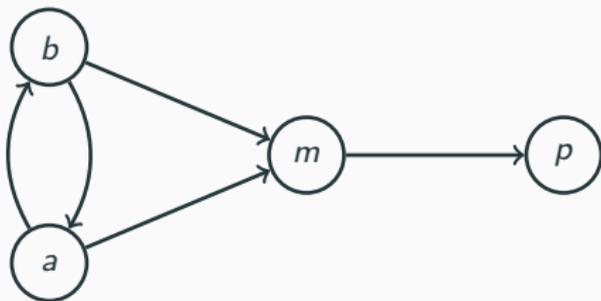
Admissible sets:  $\{\}, \{a\}, \{b\}, \{a, p\}, \{b, p\}$

Preferred, stable:  $\{a, p\}, \{b, p\}$

Grounded:

## Q2: Extracting coherent subparts of the debates

The **floating conclusion** example :



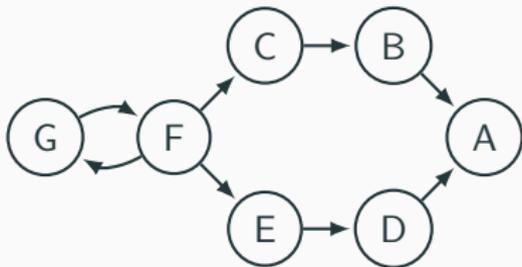
Admissible sets:  $\{\}, \{a\}, \{b\}, \{a, p\}, \{b, p\}$

Preferred, stable:  $\{a, p\}, \{b, p\}$

Grounded:  $\{\}$

### Q3: Contributions of users

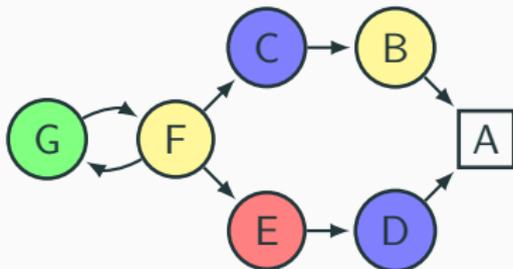
- (A) Diesel cars should be banned from in the city centre
- (B) Artisans cannot change their vehicles
- (C) The city can offer financial assistance to artisans
- (D) Autonomy of electric cars is poor, as there are not enough charging stations around
- (E) The city can set up more charging stations
- (F) The city should not spend additional money
- (G) Health and climate change issues are important



4 users: each color is a different user.

### Q3: Contributions of users

- (A) Diesel cars should be banned from in the city centre
- (B) Artisans cannot change their vehicles
- (C) The city can offer financial assistance to artisans
- (D) Autonomy of electric cars is poor, as there are not enough charging stations around
- (E) The city can set up more charging stations
- (F) The city should not spend additional money
- (G) Health and climate change issues are important



4 users: each color is a different user.

### Q3: Contributions of users

users	Cat	SAF
Y	0.500	0.009
B	0.500	0.009
R	1	0.990
G	1	0.990
YG	0.500	0.009
YB	0.375	0.000
YR	0.500	0.009
BR	0.666	0.980
BG	0.500	0.009
RG	1	0.990
YBG	0.381	0.004
YRG	0.500	0.009
YBR	0.428	0.000
BRG	0.666	0.980
YRBG	0.447	0.250

We may look at the difference in score induced by the absence of each user wrt the full debate:

Y:	-.219,	-.740
G:	+.19,	+.250
B:	-.53,	+.241
R:	+.67,	+.246

Interesting to note that:

- in absolute terms, Y is the most influential
- B has positive or negative influence, depending on the semantics
- R or G have the most positive influence, depending on the semantics

We may also look more systematically at the users' marginal contributions (as power indices).

# Conclusion

Online debate platforms raise many interesting problems!

- evaluation of arguments in (bipolar/weighted) argumentation systems
- argumentation mining techniques potentially very useful to build systems

Many more to do:

- dealing with malicious behaviours (trolls, puppets)
- game-theoretical analysis of behaviour in debates (equilibrium, best response dynamics)

Visit <http://amande.lip6.fr>