

# Debates and Democracy: Insights from abstract argumentation.

Louise Dupuis de Tarlé

LAMSADE - Université Paris Dauphine  
louise.dupuis@dauphine.eu

# Introduction: Argumentation and Democracy

# Habermas and the "discourse theory of democracy"



Jürgen Habermas (1929 -) : German Philosopher and Social Theorist.

Argumentation in the "public sphere" → Parliamentary Complex →  
Laws and Regulations

**Rules** that allow the **better argument** to triumph, and a **rational consensus** to be formed.

# Assumptions about Debates

- That the debate results in a consensus.
- That the result reflects the participant's positions.
- That the only important factors are the strength of the arguments.
- That the agents are rational in their assessment of the debate.

Exploring with models of **Abstract Argumentation** → Debates may be more complex.

# Outline

1. Introduction: Argumentation and Democracy
2. Background : Abstract Argumentation
3. Multi-Agent Models of Argumentation
4. Models of Biased Agents
5. Conclusion

## Background : Abstract Argumentation

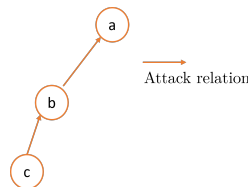
# Example of a Debate



# Abstract Argumentation Theory

Arguments are **abstract**, their content is not analysed, only the relations between them.

- a : ChatGPT is going to destroy many jobs.
- b : No, it will perform menial tasks and increase productivity.
- c : Writing articles and code are not menial tasks.



Dung, P. M.. *On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games*. Artificial intelligence 1995.

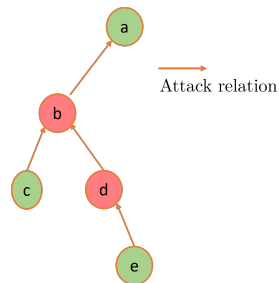


# Acceptability Semantics

Acceptability semantics are functions which return a **set of acceptable arguments**.

An argument is **acceptable** if :

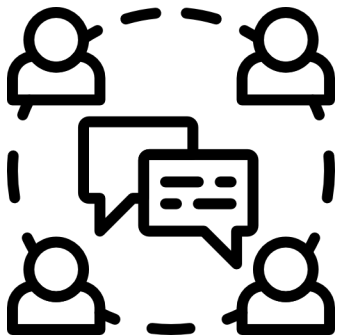
- It is not attacked.
- It is attacked only by unacceptable arguments.



Dung, Phan Minh. *On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games*. Artificial intelligence, 1995.

# Multi-Agent Models of Argumentation

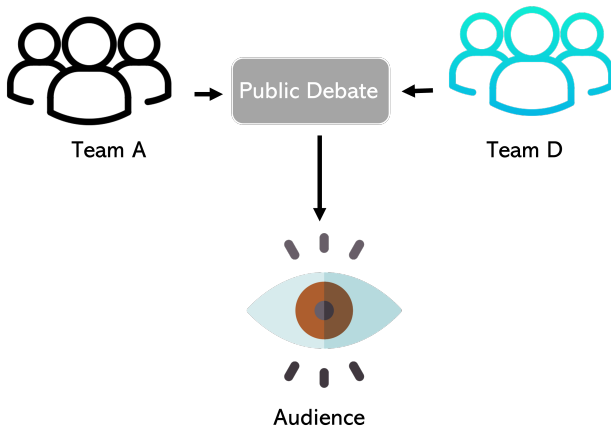
# Multi-Agent Models of Argumentation



- Agents have debate a certain **topic**
- Characterized by knowledge, behavior, strategies.

→ What is the **result** of the debate ?

# 1. Team Persuasion



Kohan Marzagão, D., Murphy, J., Young, A. P., Gauy, M. M., Luck, M., McBurney, P., Black, E. (2018). *Team Persuasion*. In Theory and Applications of Formal Argumentation: 4th International Workshop, TAFE 2017.

# 1. Team Persuasion - Example of a Bipartite Graph

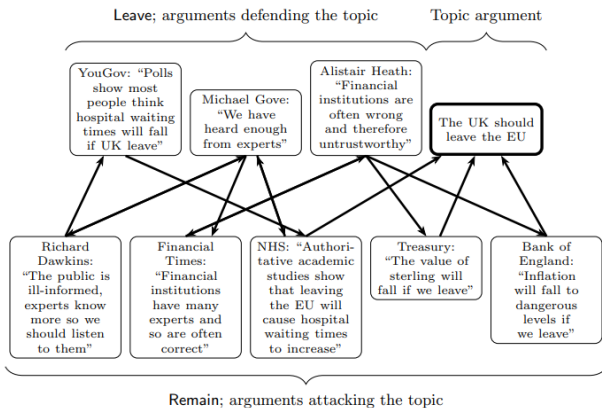


Fig. 1. An instantiated example of a bipartite argumentation framework.

# 1. Team Persuasion - Results

## State-stable configurations

A state-stable configuration is a configuration in which no agent has any reason to change their stance, and the game remains stable indefinitely.

The probability of either team winning can be computed, but there are **many games which do not become state-stable**.

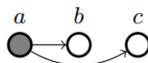
→ A debate may not yield any definitive outcome.

Temporality plays a role, e.g. in the Brexit referendum.

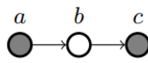
## 2. Multiparty Persuasion

Bonzon, E., Maudet, N. (2012) . *On the outcomes of multiparty persuasion*. In *Argumentation in Multi-Agent Systems: 8th International Workshop, ArgMAS 2011*.

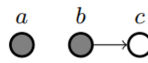
Agents are equipped with argumentation frameworks with the **same arguments** but **distinct attack relations**.



$$\mathcal{E}(AS_1) = \{a\}$$

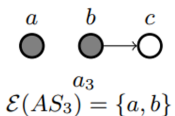
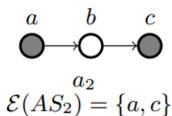
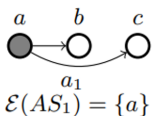


$$\mathcal{E}(AS_2) = \{a, c\}$$

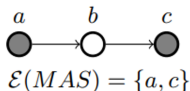


$$\mathcal{E}(AS_3) = \{a, b\}$$

## 2. Multiparty Persuasion - Merged Graph

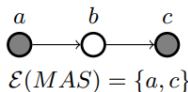
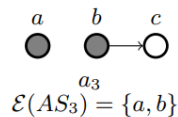
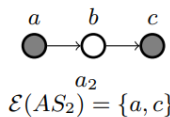
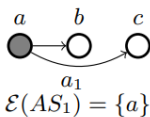


We can compute the **merged graph** by only retaining attacks supported by a (strict) majority of agents.





## 2. Multiparty Persuasion - Merged Graph



### Observation

The status of the issue in the merged argumentation system can contradict the opinion of the majority.

## 2. Multiparty Persuasion - Outcome

Agents play strategically by asserting or removing attacks from a **common gameboard**. They focus on the acceptability of a special argument, the **issue**.

*A sequence of moves allowed by the protocol is the following:*

$t = 1 - a_1$  plays for CON:  $RP_1^1 = \{(a, c)\}$

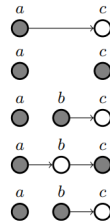
$t = 2 - a_2$ :  $RP_2^2 = \{(a, c)\}$

$t = 3 - a_3$  plays for CON:  $RP_3^3 = \{(b, c)\}$

$t = 4 - a_2$ :  $RP_2^4 = \{(a, b), (a, c)\}$

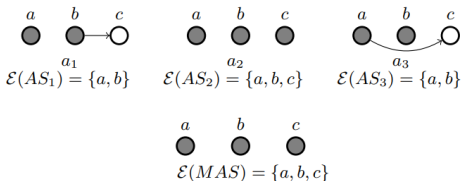
$t = 5 - a_3$ :  $RP_3^5 = \{(b, c), (a, b)\}$

$t = 6 - a_2$  cannot add  $c$  in the extension



### Observation

The outcome of a game may be different from the merged argumentation system. There may be several possible outcomes.



*The issue of the dialogue is the argument  $c$ . We have  $CON = \{a_1, a_3\}$ ,  $PRO = \{a_2\}$ . Agents in  $CON$  can attack  $c$  in two ways: either  $a_1$  can play  $bRc$ ; or  $a_3$  can play  $aRc$ . But  $a_2$  will be able to remove either attack, and  $CON$  agents will not have the possibility to counter-attack. We will obtain  $\mathcal{E}(AS(GB)) = \{a, b, c\}$ .*

## Observation

The outcome of a game may not reflect the preference of the majority of the agents.

## 2. Multiparty Persuasion

- The merged argumentation graph may not reflect the majority's preference.
- The outcome of the game is not always pre-determined.
- The outcome can be different from the merged argumentation graph.
- The outcome may not reflect the majority's preference.

How do we define the **result** of a debate ? Can debates truly represent the **preferences** of agents?



### 3. Confirmation Bias

#### Confirmation Bias

Confirmation bias is defined as people's tendency to process information by looking for, or interpreting, information that is consistent with their existing beliefs.

Dupuis de Tarlé, L., Bonzon, E., Maudet, N.. *Multiagent Dynamics of Gradual Argumentation Semantics*. . Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (pp. 363-371), 2022.

→ Confirmation bias influences the outcome of the debate, and how much the opinion of the agents converges.

## 4. MySide Bias

### MySide Bias

Introduced by Mercier and Sperber (2017), MySide bias is the tendency to **produce** arguments or reasons in favor of one's own point of view.

Not the acceptance but in the **production** of counter arguments.

Dupuis de Tarlé, L., Michelini M., Borg, A., Pigozzi G., Rouchier R., Šešelja, D., Straßer, C.. *An Agent-Based Model of MySide Bias in Scientific Debates*. Forthcoming.

- Myside bias impacts how likely the agents are at finding a correct consensus.
- Makes consensus more likely to last.

# Conclusion



# Conclusion

Even in an idealized, formal setting, collective argumentative discussions exhibit complex dynamics.

- Debates may not converge to a consensus.
- There may be several possible outcomes, and it may not represent the preferences of the agents.
- The presence of biases impacts the outcome of the debates.

→ Can we justify democracy on the premise of rational argumentative discussion ?

## Appendix

# Rules for genuine argumentation

Habermas suggests that the following rules of discourse can be established:

- ① Every subject with the competence to speak and act is allowed to take part in a discourse.
- ②
  - ① Everyone is allowed to question any assertion whatever.
  - ② Everyone is allowed to introduce any assertion whatever into the discourse.
  - ③ Everyone is allowed to express his attitudes, desires, and needs.
- ③ No speaker may be prevented, by internal or external coercion, from exercising his rights as laid down in (1) and (2)

These rules are "supposed to insulate discourse from all persuasive forces except the "unforced force of the better argument", and they must be followed, if a rationally motivated consensus is to be reached."

Finlayson, James Gordon, Dafydd Huw Rees. "*Jürgen Habermas*". The Stanford Encyclopedia of Philosophy (Winter 2023 Edition).

# Multiparty Protocol

- (1) Agents report their individual view on the issue to the central authority, which then assign (privately) each agent to PRO or CON.
- (2) The first round starts with the issue on the gameboard and the turn given to CON.
- (3) Until a group of agents cannot move, we have:
  - (a) agents independently propose moves to the central authority;
  - (b) the central authority picks the first (or at random) relevant move from the group of agents whose turn is active, update the gameboard, and passes the turn to the other group