MS dialogues: Persuading and getting persuaded

A model of social network debates that reconciles arguments and trust

Simone Gabbriellini and Paolo Torroni

University of Bologna, Italy

Background

- Model social networks debates using formal argumentation
- Rich bibliography on many related topics in argumentation
  - Argumentation in multi-agent systems
  - Persuasion dialogues
  - (and many other types of dialogues)
  - Outcomes of multi-party persuasion
  - Argumentation and trust
  - Argumentation to formalize/help/support/... online debates
  - Bottom-up argumentation
- Mainly in artifical societies settings (except last two items)
- Focus on formal properties
- Here: attempt to use argumentation to operationalize models of human interaction developed in cognitive sciences

Mercier & Sperber
Argumentative theory of reasoning

- Emergence of reasoning best understood in the framework of evolution of human communication

- Function of reasoning is argumentative

- Reasoning enables people to exchange arguments that, on the whole, make communication more reliable and hence more advantageous

- Epistemic vigilance
  - Coherence setting
  - Trust calibration

Agent reasoning

- Dung's framework
  - Arguments
  - Attacks
- Admissible extensions
- Conflict-free semantics
- May have more than one extension
- Same set of arguments, different idea of attacks

(a) sugar mills produce as much as windmills produce, and at half the cost. Therefore, sugar mills are preferable to windmills.

(b) recent studies show that windmills are much more energy-efficient than sugar mills. Therefore, windmills are preferable to sugar mills.
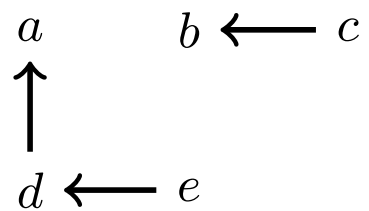
MS Dialogues: principles

- Agents use argumentative reasoning, to establish coherence of information in the posts, against their own beliefs.
- An author's input to a dialogue is coherent with the author's beliefs, i.e., it belongs to her AF or to the conclusions that can be reasonably drawn from it.
- Agents evaluate posts using mechanisms for epistemic vigilance, based on argumentation and trust.
- The trust of an agent towards another may change dynamically as the dialogue evolves.
- If a post is incoherent with the recipient agent's beliefs:
  - If the recipient trusts the post's author →belief revision to assimilate the new beliefs, while maintaining coherence;
  - If the recipient does not trust the post's author → either engage in an MS dialogue with the post's author, by producing arguments against the post, or simply ignore the post.
    - In turn, the author can produce arguments for her claims, and encourage the recipient to examine, evaluate, and accept these arguments.
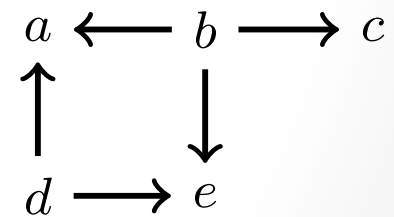   - …

MS Dialogues: an example

(a) AF_A	(b)	(c)	(d)	(e)
$a \quad b \leftarrow c$ $\uparrow$ $d \leftarrow e$	$a \leftarrow b \leftarrow c$ $\uparrow$ $d \leftarrow e$	$a \leftarrow b \leftrightarrow c$ $\uparrow$ $d \leftarrow e$	$a \leftarrow b \leftrightarrow c$ $\uparrow \qquad \downarrow$ $d \leftarrow e$	$a \leftarrow b \leftrightarrow c$ $\uparrow \qquad \downarrow$ $d \leftrightarrow e$
(f)	(g)	(h)	(i)	(j)
$a \quad b \quad c$ $\uparrow$ $d \leftarrow e$	$a \leftarrow b \quad c$ $\uparrow$ $d \leftarrow e$	$a \leftarrow b \rightarrow c$ $\uparrow$ $d \leftarrow e$	$a \leftarrow b \rightarrow c$ $\uparrow \qquad \downarrow$ $d \leftarrow e$	$a \leftarrow b \rightarrow c$ $\uparrow \qquad \downarrow$ $d \leftrightarrow e$
(k)	(l)	(m)	(n)	(o) AF_B
$a \quad b \quad c$ $\uparrow$ $d \quad e$	$a \leftarrow b \quad c$ $\uparrow$ $d \quad e$	$a \leftarrow b \rightarrow c$ $\uparrow$ $d \quad e$	$a \leftarrow b \rightarrow c$ $\uparrow \qquad \downarrow$ $d \quad e$	$a \leftarrow b \rightarrow c$ $\uparrow \qquad \downarrow$ $d \rightarrow e$

MS Dialogues: an example

(a) AF_A

$$a \qquad b \longleftarrow c$$
$$\uparrow$$
$$d \longleftarrow e$$

(o) AF_B

$$a \longleftarrow b \longrightarrow c$$
$$\uparrow \qquad\qquad \downarrow$$
$$d \longrightarrow e$$

Dialogue rules

- Initiate **[a]**
- React to initiate
  - agree: *ok*
  - trust: revise and *ok*
  - distrust: attack (and again)
- React to attack **[b→a]**
  - know attack: counter (recent or past attack)
  - don't:
    - trust: revise and reconsider (*ok* or attack)
    - don'trust: rebut
- React to rebut **[¬(b→a)]**
  - trust: revise and reconsider (*ok* or attack)
  - don't trust: *sorry*

Belief revision

- Revision of argumentation framework
  - Purpose: include information from trusted party
  - Context: human-like debate
  - Not necessarily following AGM postulates
  - Simple way: focus on attack relations between two arguments only, assuming position of counterpart
  - Conservative solution

Properties of MS Dialogues

1. MS dialogues respect agent autonomy
2. If used conservative belief revision operator, polarization does not increase
   - Conservative: no attacks artificially added or removed
3. MS dialogues stay focussed
4. Flow of dialogue is guaranteed
   - If disagreement, there exist a relevant attack
5. Agents can exhaustively express all objections to claim
6. Termination in finite number of steps
   - Upper bound is maximum number of edges

Implementation: NetArg

- NetLogo model

- Uses ConArg [Santini & Bistarelli]

- Supports many semantics, including weighted AFs

- Done some experiments

- Poster on Wednesday morning session

- Demo on Thursday afternoon session

Conclusions

- Recent research trend in argumentation
- Interdisciplinary effort to capture bottom-up nature of debates occurring in social networks
- Motivating context is social simulation
  - Effort to reach out to sociologists
  - Presentation at ESSA 2013 (http://www.essa2013.org)
- Many avenues for further research (and collaborations)
  - Evaluation using data from online debates (methodology?)
  - Organizing human authored arguments into Dung-style frameworks
  - Distance between argumentation frameworks
    - Recent work on quantifying disagreement within a single AF
  - Revision of argumentation frameworks
  - Novel integration of argumentation and trust (dialogues)

Thank you!!!

Further comments and questions:
simone.gabbriellini@unibo.it, paolo.torroni@unibo.it
AAMAS Poster & Demo