DECENTRALISED MODERATION: BRINGING DEMOCRACY TO INTERNET

Maite Lopez-Sanchez
(maite_lopez@ub.edu)
Juan Antonio Rodríguez Aguilar
(jar@iiia.csic.es)
Jordi Ganzer
Marc Serramià
MODERATION: AN INTERNET CHALLENGE

Current virtual communities: intensive human labour to deal with moderation

Policies to regulate contents:
• To decide if a content is unacceptable

Users:
• Are not aware of such regulations
• Have not been involved in their definition

BUT...

http://www.wired.com/2014/10/content-moderation/
WHAT IF PARTICIPANTS DECIDE THEIR POLICIES?

• Ostrom: “societies” that involve individuals in the definition of their rules perform better (vs. externally imposed).

• Democracy also considers people’s opinion.

➢ Our view:

Participants will be more prone to behave correctly if involved in the decision.
HOW CAN WE AGGREGATE OPINIONS?

Norm example: “Flatmates take fixed turns for dishwashing at 10 p.m.”

• **Voting:**
  • Select from different options:
    • Yes
    • No
  • Simple
  • Unjustified
HOW CAN WE AGGREGATE OPINIONS?

Norm example: “Flatmates take fixed turns for dishwashing at 10 p.m.”

• Voting:
  • Select from different options:
    • Yes
    • No
  • Simple
  • Unjustified

• Argumentation:
  • Provide arguments in favour/against:
    • $a_1$ = “10 p.m. is too late”
    • $a_2$ = “Schedule is too rigid”
    • $a_3$ = “Fair distribution”
  • Complex
  • Justified
HOW CAN WE AGGREGATE OPINIONS?

Norm example: “Flatmates take fixed turns for dishwashing at 10 p.m.”

- **Voting:**
  - Select from different options:
    - Yes
    - No
  - Simple
  - Unjustified

- **Argumentation:**
  - Provide arguments in favour/against:
    - \(a_1=\) “10 p.m. is too late”
    - \(a_2=\) “Schedule is too rigid”
    - \(a_3=\) “Fair distribution”

- **Complex**
- **Justified**
HOW CAN WE AGGREGATE OPINIONS?

Norm example: “Flatmates take fixed turns for dishwashing at 10 p.m.”

• Voting:
  • Select from different options:
    • Yes
    • No
  • Simple
  • UnJustified

• Argumentation:
  • Provide arguments in favour/against:
    • $a_1 = \text{“10 p.m. is too late”}$
    • $a_2 = \text{“Schedule is too rigid”}$
    • $a_3 = \text{“Fair distribution”}$

• Our focus:
  • Argumentation theory
  • Information fusion
Norm ex. $N$: “Flatmates take fixed turns for dishwashing at 10 p.m.”

- Arguments:
  - $a_1 =$ “10 p.m. is too late”
  - $a_2 =$ “Schedule is too rigid”
  - $a_3 =$ “Fair distribution”
Norm ex. N: “Flatmates take fixed turns for dishwashing at 10 p.m.”

• Arguments:
  • $a_1$ = “10 p.m. is too late”
  • $a_2$ = “Schedule is too rigid”
  • $a_3$ = “Fair distribution”

Should flatmates adopt this norm?
Norm ex. N: “Flatmates take fixed turns for dishwashing at 10 p.m.”

• Arguments:
  • \(a_1\) = “10 p.m. is too late”
  • \(a_2\) = “Schedule is too rigid”
  • \(a_3\) = “Fair distribution”

Should flatmates adopt this norm?

<table>
<thead>
<tr>
<th></th>
<th>(N)</th>
<th>(a_1)</th>
<th>(a_2)</th>
<th>(a_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alan gets up early 4 days/week</td>
<td>☒</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Barbara has spare time at night</td>
<td>✓</td>
<td>☒</td>
<td>☒</td>
<td>✓</td>
</tr>
<tr>
<td>Charles is keen to have a routine</td>
<td>✓</td>
<td>✓</td>
<td>☒</td>
<td>✓</td>
</tr>
</tbody>
</table>
Norm ex. N: “Flatmates take fixed turns for dishwashing at 10 p.m.”

• Arguments:
  • $a_1$ = “10 p.m. is too late”
  • $a_2$ = “Schedule is too rigid”
  • $a_3$ = “Fair distribution”

Should flatmates adopt this norm? Yes

<table>
<thead>
<tr>
<th></th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>🚦</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td>Alan gets up early 4 days/week</td>
<td>🚦</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td>Barbara has spare time at night</td>
<td>✅</td>
<td>🚦</td>
<td>🚦</td>
</tr>
<tr>
<td>Charles is keen to have a routine</td>
<td>✅</td>
<td>✅</td>
<td>🚦</td>
</tr>
</tbody>
</table>
Our contribution:

• Method based on argumentation logic to compute aggregated decisions
OPINION AGGREGATION (ARGUMENTATION LOGIC)

Our contribution:

• Method based on argumentation logic to compute aggregated decisions

Challenges:

• Easy way to gather information from all participants
• Intuitive way to show individual/collective information
HOW CAN WE AGGREGATE OPINIONS?

Norm example: “Flatmates take fixed turns for dishwashing at 10 p.m.”

• Voting:
  • Select from different options:
    • Yes
    • No

• Simple
• Unjustified

• **Argumentation:**
  • Provide arguments in favour/against:
    • $a_1 = \text{"10 p.m. is too late"}$
    • $a_2 = \text{"Schedule is too rigid"}$
    • $a_3 = \text{"Fair distribution"}$

• **Our focus:**
  • Argumentation theory
  • Information fusion
ON-LINE COMMUNITY PROTOTYPE

http://tetris.iiia.csic.es
ON-LINE COMMUNITY PROTOTYPE

- http://tetris.iii.csic.es
ON-LINE COMMUNITY PROTOTYPE

- Norm Example discussion

Participants can add arguments in favour and against a proposed norm.
HOW DO WE DECIDE IF A NORM SHOULD BE ESTABLISHED?

• Arguments in favour > arguments against
• Number of arguments? No: not all arguments should count equally
HOW DO WE DECIDE IF A NORM SHOULD BE ESTABLISHED?

• Arguments in favour > arguments against

• Ask opinions about arguments (values)
  • How we do combine “like” with dislikes values?
  • What does a neutral value mean?
HOW DO WE DECIDE IF A NORM SHOULD BE ESTABLISHED?

• Arguments in favour > arguments against

• Ask opinions about arguments (values)

• We should just consider the ones that people think are worth
  • How many people should like it?
  • How much should they like it?
ON-LINE COMMUNITY PROTOTYPE

• Our proposal

Each participant awards stars to arguments:
• 5 totally in favour
• 3 neutral
• 1 totally against
ON-LINE COMMUNITY PROTOTYPE

Our proposal

Each participant awards stars to arguments:

- 5 totally in favour
- 3 neutral
- 1 totally against

Argument rating: the farther a rating is from neutrality, the stronger its importance when computing its collective support.
ON-LINE COMMUNITY PROTOTYPE

• Our proposal

Each participant awards stars to arguments:
• 5 totally in favour
• 3 neutral
• 1 totally against

Argument rating: the farther a rating is from neutrality, the stronger its importance when computing its collective support

Argument set rating: aggregate relevant arguments
ON-LINE COMMUNITY PROTOTYPE (INFORMATION FUSION)

• Our proposal

Aggregated norm rating to decide about the norm
OPINION AGGREGATION (INFORMATION FUSION)

Our contribution:

• Information fusion and aggregation operators to combine (numerical) opinions

Challenges:

• Usability
Problem: Collective decision making about norm adoption

Contributions:

1. Argumentation theory to handle arguments in favour or against a norm
2. Information fusion and aggregation operators that combine (numerical) individual opinions
WRAP-UP

Problem: Collective decision making about norm adoption

Contributions:

1. Argumentation theory to handle arguments in favour or against a norm
2. Information fusion and aggregation operators that combine (numerical) individual opinions

Challenges:

• Social intelligence requires to deal with complexity
• Usability when dealing with social interactions, individual/collective/aggregated opinions
THANK YOU