A Vector-Based Extension of Value-Based Argumentation for Public Interest Communication

Pietro Baroni[1], *Giulio Fellin*[1], Massimiliano Giacomin[1], Carlo Proietti[2]

[1]Università degli Studi di Brescia, [2]Consiglio Nazionale delle Ricerche

2024 November 28

EPICA

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

We acknowledge financial support from MUR project PRIN 2022 EPICA “Enhancing Public Interest Communication with Argumentation” (CUP D53D23008860006) funded by the European Union - Next Generation EU.

Introduction

Public Interest Communication (PIC) aims to promote beneficial behaviours/policies through persuasive arguments.

- Challenges in PIC campaigns:
  - Ineffectiveness/backfire due to diverse audiences and poorly targeted messaging.
  - Practitioners rely on experience, lacking formal methods to analyse outcomes.
- Computational argumentation
  - Reconstruct debates: Identify arguments and their relationships (attacks/supports).
  - Assess justification: Evaluate arguments using formal semantics.
- Focus of this work: Modelling diverse audiences using vector-based value frameworks:
  - Builds on Bench-Capon's value-based approach and its extensions.
  - Attributes multi-dimensional value vectors to arguments, enabling nuanced analysis.
  - Supports theories like Schwartz's human values frameworks.

The framework

We consider a triple $\langle A, \rightarrow, A^{\text{pos}} \rangle$ where

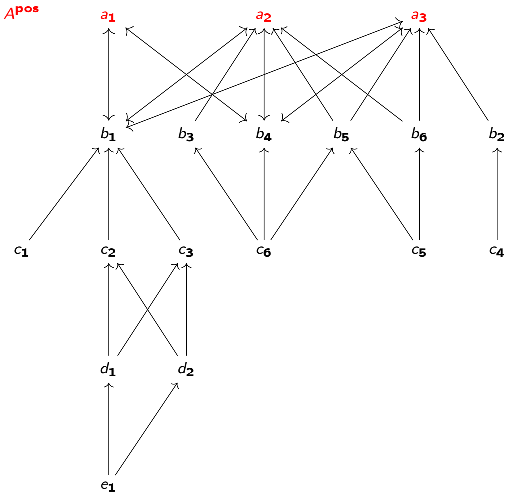
- $\langle A, \rightarrow \rangle$ is an argumentation framework, i.e.
  - A is a set of arguments, and
  - $\rightarrow$ is a binary relation $\rightarrow \subseteq A \times A$—we read $a \rightarrow b$ as "a attacks b;"
- $A^{\text{pos}} \subseteq A$ will be the set of arguments expressing the goals of the considered communication campaign.

Example—Part 1

Let's consider a campaign for a greener diet:

$$A^{\text{pos}} = \{a_1 : \text{Less chronic disease, better overall health and less foodborne illness,}$$
$$a_2 : \text{Better environment: soil, water, air,}$$
$$a_3 : \text{Less animal suffering.}\}$$

$$
\begin{aligned}
A = A^{\mathbf{pos}} \cup \{ & b_1 : \text{Veganism may be unhealthy, e.g. different blood types need different diets}, \\
& b_2 : \text{Morality is relative}, \\
& b_3 : \text{Plant-based agriculture still causes harm}, \\
& b_4 : \text{Not everyone can be vegan}, \\
& b_5 : \text{There are worse things going on in the world, this is a secondary cause}, \\
& b_6 : \text{The world is a tough place, so we have to deal with bad things}, \\
& c_1 : \text{Vegan athletes exist}, \\
& c_2 : \text{Many nutritional experts state that veganism can be healthy and optimal}, \\
& c_3 : \text{The blood-type diet theory has been debunked}, \\
& c_4 : \text{Most people are not moral relativists about unnecessary suffering}, \\
& c_5 : \text{Recognising that the world is cruel is in not an excuse to do harm}, \\
& c_6 : \text{The goal is to make progress, no one expects the world to become perfect}, \\
& d_1 : \text{Experts are influenced by financial interests and agendas}, \\
& d_2 : \text{Not all experts agree}, \\
& e_1 : \text{There is consensus among independent experts about the health benefits.} \}
\end{aligned}
$$



A^{pos}
a_1
a_2
a_3
b_1
b_3
b_4
b_5
b_6
b_2
c_1
c_2
c_3
c_6
c_5
c_4
d_1
d_2
e_1

The audiences

The set of audiences is a set of the form $I = \{1, 2, 3, ..., k\}$, of cardinality k. To each audience $i \leqslant k$ we associate a weight p_i. Weights satisfy the following conditions:

$$\forall_{i \leqslant k}\, p_i \geqslant 0, \qquad \sum_{i=1}^{k} p_i = 1.$$

The values

We define

- the space of values as $V = [0,1]^n$, each dimension of which is associated with the corresponding value;
- the value function $\mathrm{val}\colon A \to V$, which assigns each $a \in A$ to its vector of values.

Example—Part 2

In this preliminary paper, we do not argue for a specific set of values, as it falls outside the scope of the present work. For illustrative purposes we follow the list of classes of values from [Kiesel et al., 2022]:

1. Self-direction: thought
2. Self-direction: action
3. Stimulation
4. Hedonism
5. Achievement
6. Power: dominance
7. Power: resources
8. Face
9. Security: personal
10. Security: societal
11. Tradition
12. Conformity: rules
13. Conformity: interpersonal
14. Humility
15. Benevolence: caring
16. Benevolence: dependability
17. Universalism: concern
18. Universalism: nature
19. Universalism: tolerance
20. Universalism: objectivity

Example—Part 3

arg.	values																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
a_1	0	0	0	0	0	0	0	0	1	.6	0	0	0	0	0	0	0	0	0	0
a_2	0	0	0	0	0	0	0	0	0	.6	0	0	0	0	0	.7	0	1	0	0
a_3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.2	0	.6	.9	0	0
b_1	0	0	0	0	0	0	0	0	1	.6	0	0	0	0	0	0	0	0	0	0
b_2	.8	.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.2	0	.4	0
b_3	0	0	0	0	0	0	0	0	.6	0	0	0	0	0	0	.7	0	1	0	0
b_4	0	0	0	0	0	0	.9	0	0	0	0	0	0	0	0	0	.6	0	0	0
b_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.6	0	0	.7
b_6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.7
c_1	0	0	0	0	0	0	0	0	.4	0	0	0	0	0	0	0	0	0	0	.9
c_2	0	0	0	0	0	0	0	0	.6	.4	0	0	0	0	0	0	0	0	0	1
c_3	0	0	0	0	0	0	0	0	.6	.4	0	0	0	0	0	0	0	0	0	1
c_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.6	0	0	0	0	.8
c_5	0	0	0	0	0	0	0	0	0	.3	0	0	0	0	.2	.8	.6	0	0	.6
c_6	0	0	0	0	.6	0	.6	0	0	.6	0	0	0	0	0	.7	0	0	0	.6
d_1	0	0	0	0	0	.3	.4	.3	0	0	0	0	0	0	0	0	0	0	0	.7
d_2	.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.7
e_1	.4	0	0	0	0	0	0	0	.6	.4	0	0	0	0	0	0	0	0	0	.8

The audience-specific value function

Each audience $i \leqslant k$ will have their own preferences among values. We want to represent this by introducing the audience-specific value function $\mathrm{asv}\colon I \to V$, which assigns to each audience i a vector whose jth entry represents the importance that audience i gives to value j.

Example—Part 4

Suppose that $I = \{1, 2\}$, and assign to asv the following values:

i	values																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	.7	.6	.4	.3	.5	.4	.3	.4	.7	.8	.3	.4	.5	.6	.8	.7	.9	.8	.8	.7
2	.7	.8	.3	.2	.7	.6	.6	.5	.7	.6	.8	.8	.7	.5	.6	.7	.5	.5	.6	.6

We also set $p_1 = .4$ and $p_2 = .6$.

The impact measure

For audience i, the impact of an argument a is:

$$\|a\|_i = \frac{1}{\sqrt{n}} \| \mathrm{asv}(i) \odot \mathrm{val}(a) \|$$

Properties of $\| \cdot \|_i$:

- subadditivity,
- absolute homogeneity,
- monotonicity.

Example—Part 5

arg.	a_1	a_2	a_3	b_1	b_2	b_3	b_4	b_5	b_6	c_1	c_2	c_3	c_4	c_5	c_6	d_1	d_2	e_1
$\|\cdot\|_1$	.190	.236	.204	.190	.177	.236	.135	.163	.110	.154	.196	.196	.165	.208	.196	.119	.126	.183
$\|\cdot\|_2$	.176	.176	.124	.176	.186	.176	.138	.115	.094	.136	.172	.172	.134	.170	.201	.120	.113	.165

Possible Goal 1

Maximise overall effectiveness, i.e. find the $a \in A^{\text{pos}}$ such that the following quantity is maximal:

$$\sum_{i=1}^{k} p_i \cdot \|a\|_i$$

Example—Part 6

We have

$$\sum_{i=1}^{k} p_i \cdot \|a_1\|_i = 0.4\|a_1\|_1 + 0.6\|a_1\|_1 \approx 0.182;$$

$$\sum_{i=1}^{k} p_i \cdot \|a_2\|_i = 0.4\|a_2\|_1 + 0.6\|a_2\|_1 \approx 0.200;$$

$$\sum_{i=1}^{k} p_i \cdot \|a_3\|_i = 0.4\|a_3\|_1 + 0.6\|a_3\|_1 \approx 0.156.$$

Hence the chosen argument is a_2.

Possible Goal 2

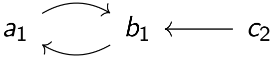
Maximise number of convinced individuals i.e. find the $a \in A^{\text{pos}}$ such that the following quantity is maximal:

$$\sum_{i=1}^{k} p_i \cdot \chi(\text{con}_i(a))$$

Proposal for convincing argument:

$$\text{con}_i(b) \iff \forall_{a \to b} \|a\|_i < \|b\|_i$$

Example—Part 7



- b_1 is not convincing as it is defeated by c_2.
- a_1 is also not convincing, as a_1 and b_1 mutually defeat.
- Intuitively, a_1 should be convincing, as its only defeater is itself defeated.

Solution: Grounded Semantics

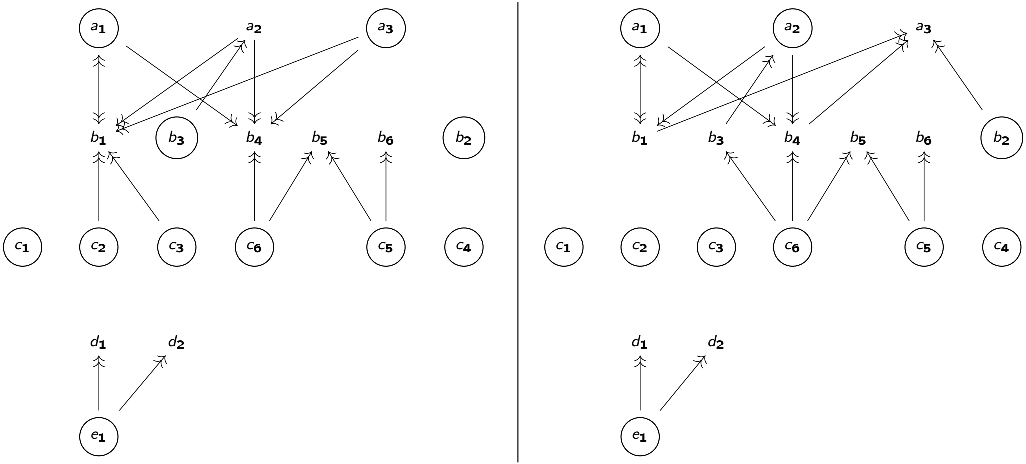
We use grounded semantics $\mathcal{E}_{GR}(A)$ to identify convincing arguments. The algorithm to compute $\mathcal{E}_{GR}(A)$:

1. Start with undefeated arguments.
2. Recursively add arguments defended by the current set.
3. Stop when no more arguments can be added.

Convincing arguments:

$$\mathrm{con}_i(a) \iff a \in \mathcal{E}_{GR}(A)$$

Example—Part 8



We observe that arguments a_1, a_3 are convincing to audience 1, while arguments a_1, a_2 are convincing to audience 2. Therefore:

$$\sum_{i=1}^{k} p_i \cdot \chi(\mathrm{con}_i(a_1)) = 0.4 \cdot 1 + 0.6 \cdot 1 = 1;$$

$$\sum_{i=1}^{k} p_i \cdot \chi(\mathrm{con}_i(a_2)) = 0.4 \cdot 0 + 0.6 \cdot 1 = 0.6;$$

$$\sum_{i=1}^{k} p_i \cdot \chi(\mathrm{con}_i(a_3)) = 0.4 \cdot 1 + 0.6 \cdot 0 = 0.4.$$

Hence the chosen argument is a_1.

Key insights and future directions

- Enhancing convincing arguments:
  - Apply Bayesian reasoning, machine learning, and datasets.
  - Tailor definitions to different campaign contexts.
- List-of-arguments campaigns:
  - Order and length of lists affect impact.
  - Ensure consistency (avoid conflicts) and cohesion (supportive arguments).
- Broader considerations:
  - Support already convinced individuals.
  - Address potential backlash or unintended effects.
  - Incorporate temporal strategies with evolving goals.
- Alternative perspectives:
  - Focus on conclusions rather than individual arguments.
  - Streamline for common ground and key divergences.

Thank you!