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

Pietro Baroni¹, Giulio Fellin^{1,*}, Massimiliano Giacomin¹ and Carlo Proietti²

¹University of Brescia, Department of Information Engineering, via Branze 38 - 25123 Brescia, Italy ²National Research Council of Italy (CNR), Institute for Computational Linguistics "A. Zampolli", Area di ricerca di Genova, Torre di Francia, Via de Marini 6 - 16149 Genova, Italy

Abstract

In this paper, we propose a mathematical model to quantify and analyse the impact of public interest communication on target audiences. Building on Bench-Capon's value-based approach, our model introduces the concept of value vectors to represent a multi-dimensional spectrum of values influencing audience perception and response. By employing vectors, we aim to capture the nuanced interplay between diverse values and the effectiveness of communication strategies.

Keywords

argumentation, value-based model, vector space

1. Introduction

Public Interest Communication plays a crucial role in promoting beneficial behaviours and policies, clarifying their rationale, and ensuring legitimacy among stakeholders, often institutions. Examples of such communication efforts include vaccination campaigns and campaigns advocating for a greener diet. Public Interest Communication aims to convince a general audience to adopt certain behaviours by presenting a variety of supporting arguments. For instance, in a campaign promoting a greener diet, arguments might include:

- Eating more fruits and vegetables benefits your health.
- Eating more fruits and vegetables potentially benefits animal welfare.
- Eating more fruits and vegetables benefits the environment.
- Eating more fruits and vegetables benefits the local economy.

Despite their importance, public interest campaigns often face significant challenges. Ineffectiveness or backfire effects, caused by poorly targeted communication, are common issues, as evidenced by numerous unsuccessful and costly campaigns. One key problem is that these campaigns target a general audience with diverse knowledge, needs, values, and attitudes. Finding a one-size-fits-all strategy is difficult, which is why campaigns typically leverage multiple motivations, as illustrated by the arguments in the greener diet example.

Moreover, public interest campaigns are primarily led by practitioners relying on experience and practical know-how. This approach makes it challenging to analyse the reasons for a campaign's success or failure. The need for such explanations has led to the emerging field of Public Interest Communication studies.

Computational argumentation offers valuable tools for Public Interest Communication by:

1. Reconstructing the general structure of a debate, identifying its basic components (arguments) and their relationships.

*AI*³ 2024 - 8th Workshop on Advances in Argumentation in Artificial Intelligence *Corresponding author.

[☆] pietro.baroni@unibs.it (P. Baroni); giulio.fellin@unibs.it (G. Fellin); massimiliano.giacomin@unibs.it (M. Giacomin); carlo.proietti@ilc.cnr.it (C. Proietti)

^{© 2024} Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

2. Providing a principled method to assess which arguments are justified in a debate, based on their status (e.g., unattacked or ultimately reinstated by other arguments) as dictated by chosen semantics.

This capability allows for an a posteriori analysis of the argumentative content and structure of public interest campaigns, potentially explaining their success or failure.

To effectively apply computational argumentation to Public Interest Communication, several challenges must be addressed:

- 1. *Argument Retrieval*: How to mine arguments from public campaigns, which requires identifying an annotation format for arguments and their mutual relations, such as attacks and supports.
- 2. *Incomplete Information*: Public interest campaign materials often provide limited information for reconstructing and evaluating the generated argumentative content. Arguments are typically from a single source (the issuing organization), with counterarguments and counter-counterarguments often implicit or from indirect sources (e.g., social forum discussions), necessitating abductive reasoning.
- 3. *Diverse Audiences*: Different individuals assess argumentative structures in various ways due to differing inferential/epistemic standards and perceptions of defeats. Additionally, real-life arguments have non-inferential components, such as promoted values, affecting their assessment.

In this preliminary paper, we place particular emphasis on addressing the challenge of *Diverse Audiences*. Our focus is on developing a mathematical model to quantify and analyse how different inferential and epistemic standards, as well as the influence of non-inferential components like promoted values, affect the assessment of public interest communication among varied target audiences. Arguments can be represented as vectors of coordinates, where each coordinate represents the strength of the argument concerning a particular value or aspect. For each individual or audience, a mapping function assigns weights to arguments, determining which arguments defeat others based on their relative weights.

Building on Bench-Capon's value-based approach [1], our model introduces the concept of value vectors to represent a multi-dimensional spectrum of values influencing audience perception and response. By employing vectors, we aim to capture the nuanced interplay between diverse values and the effectiveness of communication strategies. This approach allows for a more detailed and structured analysis of how different arguments resonate with various segments of the audience, providing insights into the reasons behind the success or failure of public interest campaigns.

In [1] each argument is attributed a single value. The hierarchy of values for a given audience determines whether or not attacks from (or to) other arguments are successful. This approach is arguably the most immediate one for representing how a value dimension may be attached to argumentative discourse. Yet, as a matter of fact, arguments may refer to more than a single value, as soon as they are more articulated. In view of this, [7] generalises Bench-Capon's approach by allowing arguments that refer to multiple values. This is implemented by a function arg attributing a (possibly empty) set of arguments to each value. Therefore, an argument a has multiple values whenever, for two distinct values v_1 and v_2 , a belongs to the intersection of $arg(v_1)$ and $arg(v_2)$; it has no value when a is not a member of arg(v) for any value v. Contrary to [1], in [7] a total ordering among values does not suffice to uniquely determine the defeat relation. To this end it is necessary to define an ordering among sets of values. In [7][Sect. 3.2] three different instances of such an ordering are introduced, deriving them from a primitive ordering among arguments.

The framework we introduce is similar in spirit to that of [7] insofar as it attributes a bundle of n features (values) to each argument, specified as a vector in an n-dimensional space. Different from [7], the vectorial representation allows for grading how much a given value attaches to one argument. This, in turn, enables a larger fan of possibilities for defining the relative weight of arguments with respect to a given audience, and therefore the success or failure of an attack. We define such a weight by means of an impact measure (Sect. 3). It is important to stress that this approach allows in principle to encompass different psychological theories of basic human values such as those by Schwartz [11] and Haidt [4].

The paper is organised as follows. Section 2 introduces the proposed formal framework, Section 3 describes the impact measure adopted, while Section 4 discusses the notion of convincing argument.

Section 5 explores a generalisation to lists of arguments included in a campaign, while Section 6 concludes the paper.

2. The framework

We propose a Value-based Argumentation Framework extending the approach in [1].

In particular, 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 $\rightarrow \subseteq A \times A$ —and $A^{\text{pos}} \subseteq A$. We read $a \rightarrow b$ as "a attacks b." The set A^{pos} will be the set of arguments expressing the goals of the considered communication campaign. In practical applications, one starts by identifying A^{pos} and then constructs A by adding possible chains of attacks to elements of A^{pos} .

By extension, we also say that a set of arguments " $S \subseteq A$ attacks b" (in symbols $S \to b$) if there is $a \in S$ such that $a \to b$. The framework $\langle A, \to \rangle$ is conveniently represented as a directed graph in which the arguments are vertices and edges represent attacks between arguments.

Example, part 1. Let's consider a campaign which aims to encourage a shift towards a greener diet by promoting the environmental and health benefits of reducing meat consumption and increasing plant-based food intake. We can identify the following set for A^{pos} :¹

 $A^{\text{pos}} = \{a_1 : \text{Less chronic disease, better overall health and less foodborne illness,} \}$

 a_2 : Better environment: soil, water, air,

 a_3 : Less animal suffering.}

To build A, we must consider possible attacks to these arguments, and attacks to them, and so on. For instance, let us take:²

 $A = A^{\text{pos}} \cup \{b_1 \colon \text{Veganism may be unhealthy, e.g. different blood types need different diets,}$

- b_2 : Morality is relative,
- b₃: Plant-based agriculture still causes harm,
- b_4 : Not everyone can be vegan,
- b_5 : There are worse things going on in the world, this is a secondary cause,
- b_6 : The world is a tough place, so we have to deal with bad things,
- c_1 : Vegan athletes exist,
- c_2 : Many nutritional experts state that veganism can be healthy and optimal,
- c_3 : The blood-type diet theory has been debunked,
- c_4 : Most people are not moral relativists about unnecessary suffering,
- c_5 : Recognizing that the world is cruel is in not an excuse to do harm,
- c_6 : The goal is to make progress, no one expects the world to become perfect,
- d_1 : Experts are influenced by financial interests and agendas,
- d_2 : Not all experts agree,
- e_1 : There is consensus among independent experts about the health benefits.}

Then $\langle A, \rightarrow, A^{\text{pos}} \rangle$ can be represented as in Figure 1.

The key question is whether a given argument $a \in A^{\text{pos}}$ will convince the public of our campaign. In order to propose some ways to answer this question, we endow $\langle A, \rightarrow, A^{\text{pos}} \rangle$ with some additional structure.

¹The list is taken from [6]. For simplicity, some arguments have been merged.

 $^{^{2}}$ We mostly follow [3]. We acknowledge that this list is not exhaustive; it is provided solely for illustrative purposes.



Figure 1: The diagram representing the framework described in Example 1.

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

$$\forall_{i \le k} \, p_i \ge 0, \qquad \qquad \sum_{i=1}^k p_i = 1$$

This takes into account the fact that each i may not represent an individual listener, but a portion of the whole population that has similar values.

Let us suppose that we have a fixed (and ordered) set of values (e.g. equality, individual health, collective health, etc.) of cardinality n.

In this preliminary paper, we do not argue for a specific set of values, as it falls outside the scope of the present work. Various solutions can be found in the literature, such as in [8, 9, 10, 11]. For illustrative purposes we follow the list of classes of values from [8]:

Example, part 2. List of values:

- 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

ara											value	S								
arg.	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

Figure 2: Tabular representation of the vectors of values of the arguments in the running example. We stress again that these values are given for illustrative purposes.

In [1], each argument is associated with an individual value. We believe that this can be refined, as an argument can rely on more than one value, each with a different degree. To this purpose, we define

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

It may be possible to refine the model by imposing certain restrictions on the vectors resulting from val. Such restrictions should be motivated by further considerations or empirical evidence.

Example, part 3. Let us consider for instance the arguments a_1, b_1, c_2 from part 1 of our running Example, i.e.

- a_1 : Less chronic disease, better overall health and less foodborne illness.
- b_1 : Veganism may be unhealthy, e.g. different blood types need different diets,
- c_2 : Many nutritional experts state that veganism can be healthy and optimal.

Argument a_1 relies highly on value 9 – 'Security: personal', which includes personal health, and to a lesser extent to value 10 – 'Security: societal'. It has little to no connection to other values. Therefore, a reasonable vector of values can be

$$\operatorname{val}(a_1) = \langle 0, 0, 0, 0, 0, 0, 0, 0, 1, 0.6, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 \rangle.$$

Notice that b_1 also relies on the very same values—although with a different conclusion—, so val (b_1) = val (a_1) . Argument c_2 instead relies solely on value 20 – 'Universalism: objectivity', hence

Figure 2 presents the vectors of values for all the arguments considered in our running example.

Each audience $i \leq k$ will have their own preferences among values. We want to represent this by introducing the *audience specific value function* asv: $I \rightarrow V$, which assigns to each audience i a vector whose jth entry represents the importance that audience i gives to value j. As above, further considerations or empirical evidence may require imposing some restrictions on the vectors resulting from asv.

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 $p_2 = .6$.

We have defined all the primitive structures that we expect the framework to have. We want to answer the question from above: given an argument $a \in A^{\text{pos}}$, will a convince the audience?

To this purpose, we introduce in next section an impact measure which aims to capture how much each argument is effective for a given audience.

3. The impact measure

For each audience $i \leq k$ we introduce their *impact measure* as a function

$$\|\cdot\|_i\colon A\to [0,1].$$

We want it to assign to each argument a the impact that a has on i, by evaluating the values on which a relies and the importance that these values have for i. In particular, we want to ensure that if an argument a relies on a value that is important to i, then a will have a high impact on i. To this end, we require that this can be written as a composition

$$\|\cdot\|_i\colon A\xrightarrow{\mathrm{val}} V\subseteq \mathbb{R}^n\xrightarrow{\mathrm{s}_i} \mathbb{R}$$

In words, given an argument a, its impact on audience i is determined on the basis of its value of vectors val(a) which is then synthesised into a single real number, measuring the impact, through an audience specific function s_i .

The function s_i is required to satisfy the following properties:

— It is a seminorm, i.e. for any $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ we have³

$$s_i(\vec{x} + \vec{y}) \le s_i(\vec{x}) + s_i(\vec{y}),$$
 (Subadditivity)

$$s_i(\lambda \cdot \vec{x}) = |\lambda| \cdot s_i(\vec{x}).$$
 (Absolute homogeneity)

- − It satisfies monotonicity: Consider $\vec{x} = \langle x_1, ..., x_n \rangle$, $\vec{y} = \langle y_1, ..., y_n \rangle$. If $|x_1| \le |y_1|, ..., |x_n| \le |y_n|$, then $s_i(\vec{x}) \le s_i(\vec{y})$.
- The restriction $s_i |_V$ has codomain contained in [0, 1].

In particular, monotonicity ensures that if an argument a relies on values which are of higher importance to i, then a will have a higher impact on i.

Seminorms satisfy some very nice properties. For instance:

Proposition 3.1. Let $s_i \colon \mathbb{R}^n \to \mathbb{R}$ be a seminorm. Then:

 $^{|| \}cdot |$ denotes the absolute value.

- 1. $s_i(\vec{0}) = 0.$
- 2. Nonnegativity: For every $\vec{x} \in \mathbb{R}^n$, we have $s_i(\vec{x}) \ge 0$.

Proof.

- 1. Absolute homogeneity implies $s_i(\vec{0}) = s_i(0 \cdot \vec{x}) = 0 \cdot s_i(\vec{x}) = 0$.
- 2. Let $\vec{x} \in \mathbb{R}^n$. Absolute homogeneity implies $s_i(-\vec{x}) = s_i((-1) \cdot \vec{x}) = |-1| \cdot s_i(\vec{x}) = s_i(\vec{x})$. Subadditivity now implies $0 = s_i(\vec{0}) = s_i(\vec{x} + (-\vec{x})) \le s_i(\vec{x}) + s_i(-\vec{x}) = s_i(\vec{x}) + s_i(\vec{x}) = 2 s_i(\vec{x})$. which, in turn, implies $0 \le s_i(\vec{x})$.

In this paper we consider the following definition for the impact measure:

$$\|\cdot\|_i \colon A \to [0,1],$$
$$a \mapsto \sqrt{\frac{1}{n} \sum_{j=1}^n \left(\operatorname{asv}(i)_j \cdot \operatorname{val}(a)_j\right)^2}.$$

If we use $\|\cdot\|$ to denote the Euclidean norm and \odot to denote the Hadamard product (i.e. the componentwise product), then we can write

$$\|a\|_{i} = \frac{1}{\sqrt{n}} \|\operatorname{asv}(i) \odot \operatorname{val}(a)\|.$$

We want to show that this function indeed satisfies the desired properties:

Proposition 3.2. The function

$$s_i \colon \mathbb{R}^n \to \mathbb{R}$$

 $\vec{x} \mapsto \frac{1}{\sqrt{n}} ||\operatorname{asv}(i) \odot \vec{x}|$

is a monotonic seminorm, and the restriction $s_i |_V$ has codomain contained in [0, 1].

Proof. First, recall that the Euclidean norm is a monotonic seminorm. In order to prove subadditivity:

$$\begin{aligned} \mathbf{s}_i(\vec{x}+\vec{y}) &= \frac{1}{\sqrt{n}} \|\operatorname{asv}(i) \odot (\vec{x}+\vec{y})\| = \frac{1}{\sqrt{n}} \|(\operatorname{asv}(i) \odot \vec{x}) + (\operatorname{asv}(i) \odot \vec{y})\| \\ &\leq \frac{1}{\sqrt{n}} \|\operatorname{asv}(i) \odot \vec{x}\| + \frac{1}{\sqrt{n}} \|\operatorname{asv}(i) \odot \vec{y}\| = \mathbf{s}_i(\vec{x}) + \mathbf{s}_i(\vec{y}). \end{aligned}$$

In order to prove absolute homogeneity:

$$s_i(\lambda \vec{x}) = \frac{1}{\sqrt{n}} \|\operatorname{asv}(i) \odot \lambda \vec{x}\| = \frac{1}{\sqrt{n}} \|\lambda(\operatorname{asv}(i) \odot \vec{x})\|$$
$$= |\lambda| \frac{1}{\sqrt{n}} \|\operatorname{asv}(i) \odot \vec{x}\| = |\lambda| s_i(\vec{x})$$

In order to prove monotonicity, consider $\vec{x} = \langle x_1, ..., x_n \rangle, \vec{y} = \langle y_1, ..., y_n \rangle$ such that $|x_1| \leq |y_1|, ..., |x_n| \leq |y_n|$. Then:

$$\mathbf{s}_i(\vec{x}) = \frac{1}{\sqrt{n}} \|\operatorname{asv}(i) \odot \vec{x}\| \le \frac{1}{\sqrt{n}} \|\operatorname{asv}(i) \odot \vec{y}\| = \mathbf{s}_i(\vec{y})$$

In order to prove that the restriction $s_i |_V$ has codomain contained in [0, 1]:

$$0 = \frac{1}{\sqrt{n}} \|\vec{0}\| \le s_i(\vec{x}) \le \frac{1}{\sqrt{n}} \|\vec{1}\| = \frac{1}{\sqrt{n}} \sqrt{n} = 1.$$

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

Figure 3: Values of the impact measures of the arguments in the example for audiences 1 and 2.

Example, part 5. In Figure 3 we approximate the values of the impact function in our running example.

Beside the one proposed here, there are several possible ways to define an impact function, and empirical evidence will be necessary to evaluate and compare these different definitions. In particular, one possibility is to consider additional features of arguments besides values, for instance their presentation form, their emotional aspects an so on. A simple way to capture these additional features would be including in the computation a multiplicative factor given by a function $E: I \times A \rightarrow [0, 1]$:

$$\|\cdot\|_i \colon A \to [0,1], \qquad a \mapsto E(i,a) \cdot \sqrt{\frac{1}{n} \sum_{j=1}^n (\operatorname{asv}(i)_j \cdot \operatorname{val}(a)_j)^2}.$$

Further investigation in this direction is left to future work.

4. Convincing arguments

In this section we explore the issue of expressing in formal terms the goals corresponding to the question: "Will the argument $b \in A$ convince the audience?"

A first simple goal that one can identify is to maximise overall effectiveness:

Goal 1 Find the argument $a \in A^{pos}$ for which 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 .

Another possible goal is to maximise the number of people convinced by an argument $a \in A^{\text{pos}}$, that is:

Goal 2 Find the argument $a \in A^{pos}$ for which the following quantity is maximal:

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

where

$$\chi(\varphi) = \begin{cases} 1 & \text{if } \varphi \text{ is true,} \\ 0 & \text{if } \varphi \text{ is false.} \end{cases}$$

and $con_i(a)$ is true iff a is able to convince the audience i.

But what does it mean for an argument to convince an audience? In other words, how do we define con_i ?

One reasonable view is that an argument b convinces the audience i if and only if every attack a on it is rejected by some condition (that may depend on b and i):

$$\operatorname{con}_i(b) \iff \forall_{a \to b} \operatorname{argument} a \text{ is rejected.}$$

How do we determine whether an argument a is rejected?

For a first tentative answer to this question, we define the following relation:

$$a \twoheadrightarrow_i b \iff (a \to b \& ||a||_i \ge ||b||_i).$$

We read $a \twoheadrightarrow_i b$ as "argument *a defeats* argument *b* according to audience *i*." We also write $a \nleftrightarrow_i b$ for $\neg(a \twoheadrightarrow_i b)$. The idea, in line with the spirit of Bench-Capon's approach, is that an argument *a* can succesfully attack another argument *b* according to a given audience *i* only if *b* is not preferred by *i* to *a* on the basis of the promoted values.

Accordingly, the notion of convincing argument can be formalized as follows::

$$\operatorname{con}_i(b) \iff \forall_{a \to b} a \not\twoheadrightarrow_i b \iff \forall_{a \to b} \|a\|_i < \|b\|_i.$$

This formalisation is rather simple and corresponds to strong notion of convincing argument: the argument is regarded as unquestionable by the audience as it does not receive any effective attack. This strong notion may however turn out to be inadequate in some cases, as discussed in the next example.

Example, part 7. Let us consider the following fragment of our framework:

$$a_1 \longrightarrow b_1 \longleftarrow c_2$$

According to the formalisation given above, argument b_1 is clearly not convincing since it is defeated by c_2 . However, a_1 is also not convincing, as a_1 and b_1 are mutually defeated. On the other hand, intuitively, we are inclined to say that a_1 is convincing since its only defeater is defeated by some undefeatable argument.

As a solution, we propose to define con_i using the grounded semantics [2], which corresponds to a recursive notion of strong defence. We recall here the basic notions of grounded semantics for the benefit of readers not already familiar with Dung's theory of argumentation frameworks. According to grounded semantics, convincing arguments can be identified as follows. We start with arguments that are not defeated by any other argument, making them unquestionable. Subsequently, arguments that are defended by these unquestionable arguments (i.e. arguments whose attackers are in turn attacked by unquestionable arguments) are also recursively considered unquestionable. More precisely, we propose the following algorithm:

- 1. Consider $\langle A, \twoheadrightarrow_i \rangle$ as an argumentation framework.
- 2. Add undefeated arguments to the grounded extension $\mathcal{E}_{GR}(A)$.
- 3. Remove from the framework A the arguments defeated by elements of $\mathcal{E}_{GR}(A)$.
- 4. If in the modified framework there are undefeated arguments, then go back to Step 2, else exit.
- 5. Define

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

It is clear that by using this definition we solve the issue encountered in Example, part 7.

Example, part 8. The outcome of the grounded semantics for audiences 1 and 2 are shown in Figures 4 and 5, respectively. 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: $\sum_{i=1}^{k} p_i \cdot \chi(\operatorname{con}_i(a_1)) = 0.4 \cdot 1 + 0.6 \cdot 1 = 1$; $\sum_{i=1}^{k} p_i \cdot \chi(\operatorname{con}_i(a_2)) = 0.4 \cdot 0 + 0.6 \cdot 1 = 0.6$; $\sum_{i=1}^{k} p_i \cdot \chi(\operatorname{con}_i(a_3)) = 0.4 \cdot 1 + 0.6 \cdot 0 = 0.4$ Hence the chosen argument is a_1 .

In addressing these considerations, future research may explore refinements or extensions of the notion of convincing argument. In particular, one could consider the application of Bayesian reasoning, machine learning techniques, or other methodologies to datasets concerning past Public Interest Communication campaigns in order to better characterize the notion of convincing argument in different contexts.



Figure 4: The diagram representing the framework $\langle A, \twoheadrightarrow_1 \rangle$ of audience 1 in Example 8. The grounded extension is emphasised by circles.

5. Possible generalisation: list-of-arguments campaigns

In several cases, campaigns consist of a list of arguments, often presented as a "top ten" or similar format. This approach aims to distil complex ideas or positions into a concise and easy to remember set of key points. This strategy aims to ensure clarity and resonance with the audience, making the message more impactful. By strategically selecting these arguments, campaigns can effectively align with the values and concerns of their audience, thereby increasing their persuasive influence. The order of these arguments can be crucial. In some contexts, the first argument holds significant weight as it captures initial attention and sets the tone for the entire message. Alternatively, the last argument can leave a lasting impression, often considered the most impactful as it resonates as the final thought with the audience. Understanding how the sequence influences perception and engagement is vital for crafting compelling campaigns that effectively communicate their intended message. Moreover, audience preferences vary regarding the length of argument lists. Some audiences may prefer a shorter list that highlights key points succinctly, allowing for easy recall and immediate impact. In contrast, others may favour a longer list that provides comprehensive coverage of different aspects of the issue, appealing to those who seek detailed information and thorough analysis.

Under these considerations, if a list $\ell \in A^*$ is given⁴, we have to consider several factors, including its *length* and its *order*. These factors can be taken care of by functions $O_i \colon A^* \to [0, 1]^*$ that satisfy $length(O_i(\ell)) = length(\ell)$.

⁴Given a set S, we denote the collection of lists of elements of S by S^* . In this respect a real interval is treated as the set of all numbers included between its endpoints.



Figure 5: The diagram representing the framework $\langle A, \twoheadrightarrow_2 \rangle$ of audience 2 in Example 8. The grounded extension is emphasised by circles.

The *j*-th element of O_i represents the effect of being in position *j* on the overall impact of an argument.

There are several other properties that we might want to consider, but we don't address them in this preliminary work. These include the *consistency* and *cohesion* of the list. Consistency refers to whether the arguments within the list attack each other, potentially leading to internal conflicts that undermine the overall persuasiveness of the list. Cohesion, on the other hand, pertains to the relationships between the arguments in the list. Cohesive arguments are those that support and reinforce each other, creating a unified and convincing narrative. Taking into account both the consistency and cohesion of the arguments can help ensure that the list is both logically sound and effectively persuasive.

For each audience $i \leq k$, we extend $\|\cdot\|_i$ to lists. We propose:

$$\|\cdot\|_i: \qquad A^* \to [0,1], \qquad \langle a_1, ..., a_m \rangle \mapsto \sqrt{\sum_{j=1}^m (O_i(\langle a_1, ..., a_m \rangle)_j \cdot \|a_j\|_i)^2}.$$

Then, as above, a first goal can be

Goal 1 Find ℓ among lists of A^{pos} of a bounded (or fixed) length, such that the following quantity is maximal:

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

Again, another possible goal is to maximise the number of people convinced

Goal 2 Find ℓ among lists of A^{pos} of a bounded (or fixed) length, such that the following quantity is maximal:

$$\sum_{i=1}^{k} p_i \cdot \chi(\operatorname{con}_i(\ell)).$$

The question is then to define a notion of convincing list.

An option is to say that a list ℓ convinces *i* if and only if at least one argument of ℓ convinces *i*. Accordingly, taking also into account the impact of the order, a possible preliminary definition is:

$$\operatorname{con}_i(\langle b_1, ..., b_m \rangle) \iff \bigvee_{j \le m} \forall_{a \to b_j} \|a\|_i < O_i(\langle b_1, ..., b_m \rangle)_j \cdot \|b_j\|_i.$$

6. Final remarks

One important question to consider is whether this approach is meaningful for representing individuals who are already convinced. These individuals have already accepted the arguments and adopted the behaviours or policies being promoted. Therefore, it is crucial to evaluate if the communication under study provides any additional value to this demographic. Specifically, we need to determine if it helps in reinforcing their beliefs, preventing backsliding, or equipping them to persuade others. Understanding the impact on already convinced individuals can help refine our strategy to maintain and strengthen their commitment.

Another critical aspect to consider is the potential side effects of a campaign. For instance, could the campaign inadvertently alienate or provoke certain groups? Might it create unintended social or psychological impacts? Can the campaign, in extreme cases, have an effect opposite to that desired? (see an example in [5]). Assessing the possible side effects is essential to ensure that the campaign does not generate negative consequences that outweigh its benefits.

We want to include a temporal dimension in future phases of this work. In doing so, it becomes crucial to consider the goals of individual steps within the campaign. Over time, the dynamics of persuasion change, and strategies must adapt accordingly. For example, if a certain number of people are already convinced, this can serve as a powerful argument in later stages of the campaign to build momentum and credibility. Highlighting the growing support can encourage others to join, leveraging social proof as a persuasive tool. Therefore, a temporal strategy should include phased goals and tailored messages that evolve with the campaign's progress.

It is also important to explore alternative approaches to our current strategy. One such approach could involve comparing conclusions rather than arguments. Instead of focusing on the individual arguments and their interrelations, we could evaluate the overall conclusions reached by different groups or individuals. This method might provide a clearer understanding of the end results and help identify common ground or major points of divergence. By comparing conclusions, we can potentially streamline the analysis and focus on the most impactful elements of the debate, leading to more effective communication strategies.

Acknowledgments

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.

References

 Bench-Capon, T.J.M. (2002). Persuasion in Practical Argument Using Value-based Argumentation Frameworks. Journal of Logic and Computation. 13. 10.1093/logcom/13.3.429.

- [2] Dung, P.M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence, 77:321–357.
- [3] Elwood, Z. (2020). The best, most logical arguments against veganism. *Medium*. https: //apokerplayer.medium.com/the-best-most-logical-anti-vegan-arguments-477ebcc8aee1 (consulted on July 30, 2024)
- [4] Haidt, J., Joseph, C. (2004). Intuitive ethics: how innately prepared intuitions generate culturally variable virtues. Daedalus. 133(4):55–66
- [5] Hornik, R., Jacobsohn, L., Orwin, R., Piesse, A., Kalton, G. Effects of the National Youth Anti-Drug Media Campaign on youths. Am J Public Health. 2008 Dec;98(12):2229-36. doi: 10.2105/A-JPH.2007.125849. Epub 2008 Oct 15. PMID: 18923126; PMCID: PMC2636541.
- [6] Jacobson, M. (2006). *Six Arguments for a Greener Diet. How a More Plant-Based Diet Could Save Your Health and the Environment.* Center for Science in the Public Interest, Washington.
- [7] Kaci, S., van der Torre, L. (2008). Preference-based argumentation: Arguments supporting multiple values. International Journal of Approximate Reasoning. 48(3):730–751.
- [8] Kiesel, J., Alshomary, M., Handke, N., Cai, X., Wachsmuth, H., Stein, B. (2022). Identifying the Human Values behind Arguments. 4459–4471. https://www.doi.org/10.18653/v1/2022.acl-long. 306.
- [9] van der Meer, M., Vossen, P., Jonker, C., Murukannaiah, P. (2023). Do Differences in Values Influence Disagreements in Online Discussions?. 15986-16008. https://www.doi.org/10.18653/v1/ 2023.emnlp-main.992.
- [10] Qiu, L., Zhao, Y., Li, J., Lu, P., Peng, B., Gao, J., Zhu, S. (2022). ValueNet: A New Dataset for Human Value Driven Dialogue System. *Proceedings of the AAAI Conference on Artificial Intelligence*. 36. 11183-11191. https://www.doi.org/10.1609/aaai.v36i10.21368.
- [11] Schwartz, S.H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lönnqvist, J.-E., Demirutku, K. et al. (2012). Refining the theory of basic individual values. *Journal of personality and social psychology*, 103(4). https://www.doi.org/10.1037/a0029393.