A NetLogo Tool for Exploring Value-Based Argumentation in Public Interest Communication

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

[1]University of Brescia, Department of Information Engineering, via Branze 38 - 25123 Brescia, Italy

[2]National 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

We present a NetLogo-based tool for simulating how public interest arguments influence diverse audiences over time. Extending a previous theoretical model, agents are assigned profiles represented by value vectors that evolve through interaction with neighbours, capturing social influence dynamics. The tool computes the variation over time of the persuasive impact of arguments on the population on the basis of these evolving profiles. While the model simplifies argument exposure as continuous and uniform, it offers a foundation for more realistic simulations incorporating multiple arguments and competing campaigns in future work.

Keywords

Computational Argumentation, Public Interest Communication, Vector-Based Models, Value-Based Argumentation, NetLogo, Voting Models

1. Introduction

Motivation: Public Interest Communication Public Interest Communication is crucial for promoting beneficial behaviours and policies by clarifying their rationale and ensuring legitimacy among stakeholders, often institutions. Examples include vaccination campaigns and initiatives advocating for greener diets. Such campaigns present multiple supporting arguments, e.g., promoting fruit and vegetable consumption for health, animal welfare, environmental, and economic benefits [1]. Despite their importance, public interest campaigns often suffer from ineffectiveness or backfire effects due to the challenge of addressing diverse audiences with different knowledge, values, and attitudes. The difficulty of finding a one-size-fits-all approach motivates the use of multi-faceted argumentation strategies. Computational argumentation provides a tool to reconstruct the structure of a debate and assess which arguments are justified, and therefore allows for a posteriori analyses explaining the campaign outcomes. However, several challenges remain, among which how to address diverse audience perceptions.

In this work, we develop a multi-agent simulation model to address this challenge. The model operationalises and extends the theoretical analysis introduced in [1]. While the earlier framework provides a static, conceptual analysis of audience-specific argument evaluation, here we build on its principles to implement a dynamic model where audiences interact and influence each other over time. This simulation aims to explore how value-based heterogeneity and social influence mechanisms shape the overall impact of public interest communication efforts.

Inspiration: the NetLogo voting model The model presented in this paper draws methodological inspiration from the NetLogo Voting Model by Wilensky [10], a simple yet effective cellular automaton simulating the evolution of voting preferences within a spatially distributed population. In Wilensky's model, voters are spatially distributed on a bidimensional grid and each individual voter updates its preference based on the majority opinion among its eight immediate neighbours (the Moore neighbourhood). Additional rules allow tie-breaking or favouring the minority opinion in closely

AI[3] *2025: 9th Workshop on Advances in Argumentation in Artificial Intelligence, September 13, 2025, Rende, Italy*

*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)

contested situations, producing various collective dynamics. We extend this basic idea by moving from binary opinions (e.g., "blue" or "green" votes) to multi-dimensional value profiles. In our model, each patch represents an individual endowed with a vector of values that determine their predisposition toward certain arguments. Through repeated interactions, individuals adjust their value vectors based on the influence of their neighbours, simulating a social adaptation process. This richer representation enables the exploration of how the persuasive impact of an argument evolves over time as the distribution of audience values changes due to social interaction.

More in general the model we are developing falls in the recent tradition of argument-based models of opinion dynamics [6, 5, 9, 7, 3].

2. Model description

This computational tool aims to explore how the perceived strength of a given argument evolves within a population as individuals influence one another through social interaction. Specifically, it addresses the question:

> *Given a certain argument, how does the audience's perception of it change through interaction with neighbours?*

The theoretical framework proposed in [1] provided a static evaluation of argument strength across predefined audience types. While that framework was useful for assessing how argument reception and acceptance is influenced by the recipient's value profile, it cannot account for the dynamic effects of social influence—that is, how individuals might adapt their evaluative stance in response to their peers. To address this gap, we simulate a population of agents-represented as patches on a grid—each endowed with a vector of value preferences (e.g., health, environmental concern, economic priorities). The simulation tracks how these individual profiles evolve through repeated local interactions and how this, in turn, affects the overall persuasive impact of the argument over time. The model captures two key dynamics: (i) the adaptation of individual value profiles via neighbourhood interaction, and (ii) the real-time computation of the persuasive impact of a given argument on the population as a whole. This section details the model components and update mechanisms.

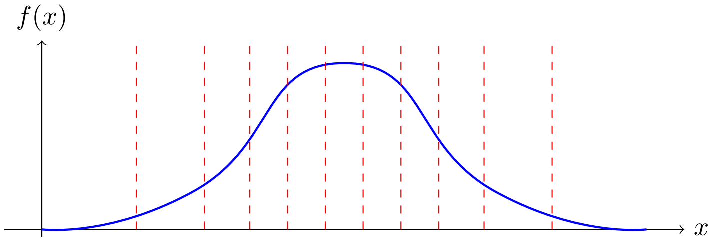
Initial setup At initialization, each patch is assigned a vector of real-valued components representing the individual's preferences over n distinct value dimensions. These values are drawn uniformly at random from the discrete set $\{0.0, 0.1, 0.2, \ldots, 1.0\}$ to reflect the heterogeneity of the population. The argument presented to the population is also described by a fixed n-dimensional vector, representing the strength of the argument along each value dimension. In the example, we chose $n = 20$ dimensions, following the number of values listed in S. Schwartz' theory of basic human values [8]. In the current version of the model, this argument vector is assumed to be uniformly and constantly available to all agents during the simulation.

Update procedure ("go" routine) At each simulation tick (representing a discrete time step), each patch re-evaluates its n-dimensional value vector by considering the corresponding components of its eight immediate neighbours (the Moore neighbourhood). This models the idea that individuals gradually adapt their values through local social influence. For each component of the value vector, the update mechanism proceeds as follows:

- **Neighbourhood summation**: For each component of the value vector, the patch computes the sum of that component across its eight Moore neighbours. Since each value lies within the interval $[0.0, 1.0]$, the resulting sum ranges from 0.0 to 8.0.
- **Normal interpretation**: This sum is interpreted using a normal distribution with mean 4.0 (the expected average of the sum of eight uniformly random values from $[0.0, 1.0]$) and variance 4

range	gets closer to
$[0.0, 1.2]$	0.0
$[1.3, 2.1]$	0.1
$[2.2, 2.7]$	0.2
$[2.8, 3.2]$	0.3
$[3.3, 3.7]$	0.4
$[3.8, 4.2]$	0.5
$[4.3, 4.7]$	0.6
$[4.8, 5.2]$	0.7
$[5.3, 5.8]$	0.8
$[5.9, 6.7]$	0.9
$[6.8, 8.0]$	1.0

(a) Mapping of neighbourhood sum ranges to target values.



(b) Normal distribution split into 11 equal-area intervals.

Figure 1: Update rule based on neighbourhood sum: value components are adjusted depending on which interval their local sum falls into. Left: the ranges and associated convergence values. Right: the conceptual partitioning of a normal distribution with mean 4.0 and variance 4.

(empirically chosen to produce reasonable spread). As shown in Figure 1b, the interval $[0.0, 8.0]$ is divided into 11 subranges of approximately equal probability mass under this distribution.[1]

- **Adjustment rule**: The patch updates each component by following these steps: (i) identify the new target value corresponding to the interval (see Table in Figure 1a); (ii) compute the average between the current value and this target; (iii) discretise the result by rounding it slightly toward the original value; (iv) assign the final value. This process is summarised in Algorithm 1.
- **Termination**: If at least one value component differs from its previous state, the patch is marked as changed. If no patch in the grid changes during a tick, the simulation halts. Alternatively, the user can explicitly decide when to terminate the simulation based on specific criteria.

Recolouring and impact calculation In parallel with the value update, the model computes the persuasive impact of the argument on each patch. This is done by taking into account the patch's value vector and the argument vector.

The impact for a patch is calculated using a weighted Euclidean norm:

$$\text{impact} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(v_i \cdot w_i)^2}$$

where:

- v_i is the i-th component of the patch's value vector,
- w_i is the i-th component of the argument vector,
- $n = 20$ is the number of dimensions.

This impact value influences the patch's colour in the visualisation, representing how strongly the argument affects each agent. See Figure 2 for an example of visualising argument impact before and after simulation. White indicates neutrality (impact $= 0.5$), green shades indicate increasing agreement (> 0.5), and red shades indicate increasing disagreement (< 0.5). It is interesting to note that social interactions lead to the formation of islands with similar opinions starting from a completely random initial state. More detailed simulations and systematic analysis are left for future work.

[1]While this approach is a simplification, it offers a basis for refinement in future work, e.g., using more principled statistical models via tools like R integrated with NetLogo.

Algorithm 1: Patch Value Update (Pseudocode)

```
foreach patch do
    foreach component i from 1 to 20 do
        total ← sum of i-th component in 8 neighbours
        previous ← current value of i-th component
        if total ∈ [0.0, 1.2] then
            target ← 0.0
        else
            if total ∈ [1.3, 2.1] then
                target ← 0.1
            else
                ...(rest of ranges omitted for brevity)
            end
        end
        new_value ← average(previous, target)
        if new_value ≠ previous then
            mark patch as changed
            update component i with new_value
        end
    end
end
```

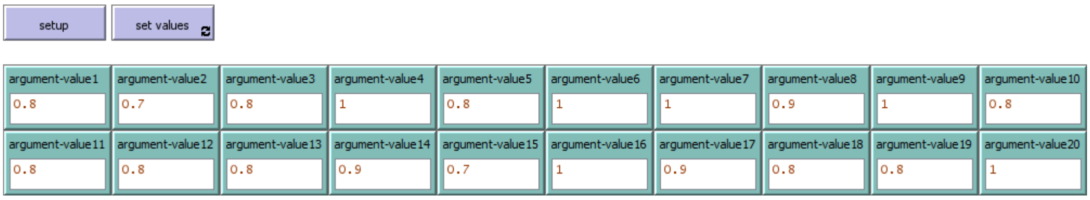
3. Conclusions and future research

In this paper, we presented a prototypical computational tool aimed at simulating audience interaction dynamics in public interest communication campaigns. Building on the value-based argumentation framework proposed in [1], this model extends the theoretical account by introducing an agent-based simulation where audience members (represented as patches) adjust their personal value profiles through local interactions.

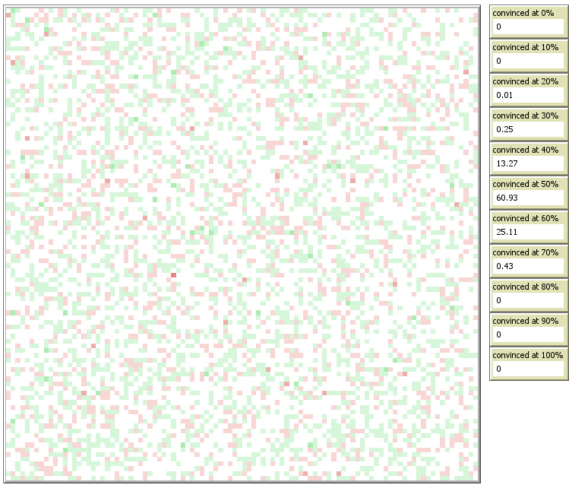
A key simplification in the current version is the assumption that the argument is constantly and uniformly presented to all agents. While this provides a controlled setting to study value adaptation and initial susceptibility, it does not reflect the uneven and selective exposure typical of real-world campaigns. Similarly, the decision rule for value updating—based on comparing the neighbourhood sum with fixed intervals derived from a normal distribution (mean 4.0, variance 4)—offers a simple approximation of social influence. Future refinements could adopt more principled statistical modelling, potentially using external tools such as R.

Despite these limitations, the present model serves as a useful starting point for investigating how audience interaction shapes the effectiveness of argumentation strategies. It complements the previous framework by operationalising value adaptation and generating structured population profiles.

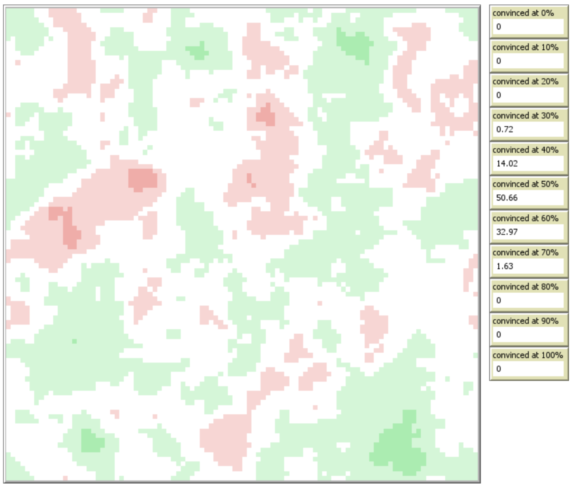
Future developments include more realistic modelling of argument dissemination (temporal, spatial, and selective exposure), and support for *multiple arguments*, allowing the simulation of campaigns that deploy varied persuasive points for different audience segments. We also plan to incorporate *opposing arguments*, either raised by the audience or from counter-campaigns, to simulate contested environments. Additional work includes refining argument retrieval, handling incomplete information, and validating the model with real campaign data. Also, exploring the effects of the social network topology on the outcome of a campaign is an interesting related issue, since this has been shown to be a very influential factor on dynamics of opinion diffusion [4].



(a) Initial setup with argument configuration and control buttons.



(b) Persuasion distribution before simulation.



(c) Persuasion distribution after simulation.

Figure 2: Example run of the simulation. (a) The initial setup where users configure the argument's value profile and start the simulation. (b) Distribution of persuasion values across agents before interaction. (c) Distribution after neighbourhood-based updates. **Colour scale:** white indicates neutrality (impact $= 0.5$); green shades represent favourable responses (> 0.5), and red shades indicate opposition (< 0.5), with darker hues signifying stronger sentiment.

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, mission 4, component 2, investment 1.1.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT-4o for grammar and spelling check. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] Pietro Baroni, Giulio Fellin, Massimiliano Giacomin, and Carlo Proietti. A vector-based extension of value-based argumentation for public interest communication. In Carlo Proietti and Carlo Taticchi (eds.), *Advances in Argumentation in Artificial Intelligence 2024*, volume 3871 of CEUR Workshop Proceedings, December 2024.

[2] Trevor J. M. Bench-Capon. Persuasion in Practical Argument Using Value-based Argumentation Frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003. https://doi.org/10.1093/logcom/13.3.429.

[3] Louise Dupuis De Tarlé, Matteo Michelini, Anne Marie Borg, Gabriella Pigozzi, Juliette Rouchier,

Dunja Šešelja and Christian Straßer. An agent-based model of myside bias in scientific debates. *Journal of Artificial Societies and Social Simulation*, 27(3), 2024. https://www.jasss.org/27/3/1.html.

[4] Andreas Flache and Michael W. Macy. Local convergence and global diversity: From interpersonal to social influence. *Journal of Conflict Resolution*, 55(6):970–995, 2011. doi:10.1177/0022002711414371.

[5] Simone Gabriellini and Paolo Torroni. A New Framework for ABMs Based on Argumentative Reasoning. In Kamiński, B., Koloch, G. (eds) *Advances in Social Simulation*. Advances in Intelligent Systems and Computing 229:25–36, 2014. https://doi.org/10.1007/978-3-642-39829-2_3.

[6] Michael Mäs and Andreas Flache. Differentiation without distancing. Explaining bi-polarization of opinions without negative influence. *PloS one*, 8(11):e74516., 2013. https://doi.org/10.1371/journal.pone.0074516.

[7] Carlo Proietti and Davide Chiarella. The role of argument strength and informational biases in polarization and bi-polarization effects. *Journal of Artificial Societies and Social Simulation*, 26(2), 2023. http://jasss.soc.surrey.ac.uk/26/2/5.html.

[8] Schwartz, S.H. et al. Refining the theory of basic individual values. *Journal of Personality and Social Psychology*, 103(4), 663–688, 2012. https://doi.org/10.1037/a0029393.

[9] Daniel Singer, Aaron Bramson, Patrick Grim, Bennett Holman, Jiin Jung, Karen Kovaka, Anika Ranginani and William J. Berger. *Rational social and political polarization*, 176(9), 2019. https://doi.org/10.1007/s11098-018-1124-5.

[10] Uri Wilensky. NetLogo Voting Model. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, 1998. http://ccl.northwestern.edu/netlogo/models/Voting