A Logical Approach to Extreme Opinion Diffusion

Enzo Battistella and Laurence Cholvy

ONERA, 2 avenue Edouard Belin, 31055 Toulouse, France

Keywords: Opinion Diffusion, Logic, Modelisation, Simulation.

Abstract: This paper focuses on diffusion of extreme opinions among agents which influence each other. In this work, opinions are modeled as formulas of the propositional logic or equivalently, as sets of propositional interpretations. Moreover, we assume that any agent changes its opinion by merging the opinions of its influencers, from most to least influential. We propose a first definition of extreme opinions and extremism. We then consider degrees of extremism. Formal studies of these definitions are made as well as simulations.

1 INTRODUCTION

Understanding the dynamics of the diffusion of opinions and especially of extremism is a tremendous question in Multi-Agent System and Artificial Intelligence communities. See for instance (Crawford et al., 2013; Christoff and Hansen, 2015; Cholvy, 2016; Grandi et al., 2015; Hafizoglu and Sen, 2012; Jager and Ambillard, 2004; Tsang and Larson, 2014; Christoff and Grossi, 2017) for the study of opinion diffusion and (Chau et al., 2014; Deffuant et al., 2002; Sureda et al., 2017) for the study of extremism diffusion.

Opinions are usually represented by a single real value between 0 (or -1) and 1 corresponding to the adequacy to a given affirmation: the closer to 1 an agent’s opinion is, the more this agent agrees with the affirmation. However some works, mainly in Artificial Intelligence community, represent an opinion by a single binary vector (Grandi et al., 2015; Christoff and Grossi, 2017) corresponding to adequacies to several affirmations. For instance, if the two affirmations are “Canada will host Winter Olympics in 2026” and “there will be acroski trials”, then the vector (1, 0) represents the opinion of an agent which thinks that Canada will host Winter Olympics in 2026 but that there will not be acroski trials.

In the present paper, we consider the model of Influence-Based Opinion Diffusion Structure (IODS) defined in (Cholvy, 2016) in which opinions are represented by propositional logic formulas or, equivalently, sets of binary vectors. There, the set \{1, 1, 0, 1, 0, 0\} represents the opinion of an agent which thinks that, if the hosting country is Canada then there will be acroski trials.

As for extremism, when opinions are represented by a single real value, it is obviously defined by having an opinion which is close to 0 (or -1) or to 1. Moreover, in such models, agents can easily be classified from most to least extremist. But for IODS, this is not so obvious and definitions of extreme opinions and extremists have to be proposed. This is the aim of the present paper.

A difficulty we will face is to define structures of influence that are relevant in some real cases. This difficult task, aiming at characterizing social networks, has been paid a huge attention to. One of the initiators of this domain was Milgram (Milgram, 1967) who proposed the theory of six degrees of separation or “small world” phenomenon. In response to this theory, several authors in graph theory defined models to describe this kind of networks (Easley and Kleinberg, 2010; Prettejohn et al., 2002; Watts and Strogatz, 1998; Watts, 1999). The difficulty dwells on adapting these models to IODS.

This paper is organized as follows. Section 2 recalls the definition of IODS. In section 3, we list some properties of extremism diffusion inspired from sociology. In section 4, we define extremism so that extremists are people with rather strong opinions. Then, we study some properties about this definition. In section 5, we introduce degrees of extremism and we study some properties of this definition. In section 6, we present and comment several simulations. Finally, section 7 presents some perspectives of this work.
2 THE UNDERLYING DIFFUSION MODEL

We consider a finite propositional language $L$. The set of its interpretations is denoted $\mathcal{I}(L)$. A multi-set of formulas $\{\varphi_1, \ldots, \varphi_n\}$ equipped with a total order $\prec$ s.t. $\varphi_i \prec \varphi_{i+1}$ ($i = 1, \ldots, n-1$) is called an ordered multi-set of formulas and denoted $\varphi_1 \prec \ldots \prec \varphi_n$.

**Definition 1** (Importance-Based Merging Operator),
An Importance-Based Merging Operator is a function $\Delta$ which associates a formula $\mu$ and a non-empty ordered multi-set of consistent formulas $\varphi_1 \prec \ldots \prec \varphi_n$ with a formula denoted $\Delta_\mu(\varphi_1 \prec \ldots \prec \varphi_n) = \min_{\varphi \in \mu} \{\varphi \prec \varphi_n\}$.

$\Delta$ is the finite set of agents. The formula $\mu$ represents the information which is true in the world. It is called integrity constraint. For any agent $i$, the formula $B_i$ represents its initial opinion. We assume that agents are rational and thus that $B_i$ is consistent and satisfies the integrity constraint $\mu$. $\text{Inf}(\varphi)$ is the non-empty set of agents which influence agent $i$. These influencers are ranked so that $j \prec i$ means that, according to $i$, its own opinion is more influenced by $j$'s opinion than by $k$'s opinion.

**Definition 3** (Influence-Based Opinion Sequence),
Let $S = (A, \mu, B, \text{Inf})$ be an IODS and $i \in A$ with $\text{Inf}(i) = \{i_1 \prec i_2 \prec \ldots \prec i_n\}$. The Influence-Based Opinion Sequence of $i$, denoted $\left( B_i^s \right)_{s \in \mathbb{N}}$, is defined by:

(i) $B_i^0 = B_i$

(ii) $\forall s > 0, B_i^s = \Delta_\mu(B_i^{s-1} \prec \ldots \prec B_i^{s-1})$

The Influence-Based Opinion Sequence (or Opinion Sequence for short) of agent $i$, $\left( B_i^s \right)_{s \in \mathbb{N}}$, represents the history of $i$'s opinion evolution. This evolution is done according to operator $\Delta_\mu$; $i$'s opinion at step $s$ is the result of applying the ordered multi-set of its influencers' opinions: $B_i^{s-1} \prec \ldots \prec B_i^{s-1}$.

**Example 1.** Consider a language with propositional letters $a$ and $b$. Let $S = (A, \mu, B, \text{Inf})$ be an IODS with: $A = \{1, 2, 3\}$, $\mu$ is a tautology, $B_1 = \neg a$, $B_2 = a \lor b$, $B_3 = \neg b$, $\text{Inf}_{f_1} = \{1\}$, $\text{Inf}_{f_2} = \{2 \prec 3\}$, $\text{Inf}_{f_3} = \{3 \prec 2\}$. Table 1 shows the evolution of the agents opinions.

<table>
<thead>
<tr>
<th>$s$</th>
<th>$i_1 = 1$</th>
<th>$i_2 = 2$</th>
<th>$i_3 = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s = 0$</td>
<td>$\neg a$</td>
<td>$\neg a$</td>
<td>$\neg a$</td>
</tr>
<tr>
<td>$s = 1$</td>
<td>$\neg a \land \neg b$</td>
<td>$\neg a \land \neg b$</td>
<td></td>
</tr>
<tr>
<td>$s \geq 2$</td>
<td>$\neg a \land \neg b$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Finally, we recall three definitions which will be useful in the rest of the paper.

**Definition 4** (Self-confident agent), Let $S = (A, \mu, B, \text{Inf})$ and $i_1 \in A$. $i_1$ is self-confident iff $\text{Inf}(i_1) = \{i_1 \prec i_2 \prec \ldots \prec i_n\}$ with $n_i \geq 1$.

**Definition 5** (Dogmatic agent). Let $S = (A, \mu, B, \text{Inf})$ and $i \in A$. $i$ is dogmatic iff $\text{Inf}(i) = \{i\}$.

**Definition 6** (Sphere of Influence of an Agent). Let $S = (A, \mu, B, \text{Inf})$ and $i \in A$. The sphere of influence of $i$ is defined by: $\text{Sphere}(i) = \bigcup_{k \geq 1} \text{Sphere}^k(i)$ with

- $\text{Sphere}^1(i) = \{j_k : \text{Inf}(j_k) = \{i \prec \ldots \prec i\}\}$
- $\text{Sphere}^k(i) = \{j_k : \text{Inf}(j_k) = \{j_{k-1} \prec \ldots \prec i\}\}$ and $j_{k-1} \in \text{ Sphere}^{k-1}(i)$.
3 PROPERTIES OF EXTREMISM DIFFUSION

In this section, we recall some observations made in the sociology literature (Bronner, 2009), (Atran, 2015) about extremism diffusion and we translate them in our setting as properties. Notice that these observations are not supposed to be general rules but only tendencies. In the two following sections, we will check if these properties are satisfied or not by the definitions we will introduce.

The first observation about extremism is that in general extremist people are close-minded, sure of their opinions and listen exclusively people with similar beliefs. But, on the other hand, extremism can and often arise in groups of friends with moderate opinions. Atran highlighted the fact that extremist people have often been introduced to extremism by a social connection.

We have retained the following five properties for our study. The first property states that if all the influencers of an agent are extremist, then after updating its opinion, the agent is extremist. The second states that if all the influencers of an agent are moderate (i.e. not extremist), then after updating its opinion the agent is moderate. The third one states that extremist agents are more certain of their opinions than moderate ones. The fourth states that extremist agents are more influential in the sense that an agent influenced by extremist agents should be "more prone" to become extremist than moderate. The fifth property, called persistence of discredited beliefs, states that an opinion will persist even if the evidence that made the opinion emerge is removed.

S being the IODS \((A, \mu, B, Inf)\) and \(N^*\) being the set of integers strictly greater than 0, these properties are formalized as follows:

**Definition 7.**

- **(P1)** \(S\) satisfies (P1) iff \(\forall i \in A, \forall t \in N, \text{if } \forall j \in Inf(i), j \text{ is extremist at step } t, \text{ then } i \text{ is extremist at step } t + 1.\)
- **(P2)** \(S\) satisfies (P2) iff \(\forall i \in A, \forall t \in N, \text{if } \forall j \in Inf(i), j \text{ is moderate at step } t, \text{ then } i \text{ is moderate at step } t + 1.\)
- **(P3)** \(S\) satisfies (P3) iff \(\forall i \in A, \forall j \in A, \forall t \in N, \text{if } i \text{ is extremist at } t \text{ and } j \text{ is moderate at } t \text{ then } \|\text{Mod}(B_i^t)\| < \|\text{Mod}(B_j^t)\|\).  
- **(P4)** \(S\) satisfies (P4) iff \(\forall t \in N, \text{there are more agents } i \text{ which are extremist at step } t + 1 \text{ and which main influencers were extremist at step } t \) than agents \(i\) which are moderate at step \(t + 1\) and which main influencers were extremist at step \(t\).
- **(P5)** \(S\) satisfies (P5) iff \(\forall t \in N, \forall i \in A, \text{if } i \text{ is extremist at step } t + 1, \text{ its main influencer was extremist at step } t \text{ and moderate at step } t + 1, \text{ then } i \text{ remains extremist at step } t + 2.\)

The previous definitions are applicable only if we have a binary definition of extremism stating whether a given agent is extremist or not. In case we have a definition stating only if an agent is more extremist than an other agent \(j\) at some step \(t\), denoted \(d_{\text{extr}}(i) \geq d_{\text{extr}}(j)\), we adapt the properties as follows:

The first property states that if all the influencers of an agent are very extremist at some step, then after the opinion updating the agent should be very extremist. The second one says symmetrically that if all the influencers of an agent are very moderate, then after the opinion updating the agent should be very moderate. The third expresses the fact that an agent should be more certain than less extremist ones in the sense of an agent’s opinion should have less models than the ones of less extremist agents. The fourth states that the extremist agents should be more influential in the sense that any agent influenced by more extremist agents should become more extremist. The fifth property, called persistence of discredited beliefs, states that an opinion will persist even if the evidence that made the opinion emerge is removed. At these properties we can add two reality constraints stating that an agent cannot have its degree of extremism increase "too much" above (respectively decrease "too much" below) the maximal degree of extremism of its influencers.

\(L_{\text{extr}} \in N^*\). \(S\) satisfies (Q1) iff \(\forall i \in A, \forall t \in N, \text{if } \forall j \in Inf(i), d_{\text{extr}}(j) \geq L_{\text{extr}}, \text{then } d_{\text{extr}}(i) \geq L_{\text{extr}}.\)

\(L_{\text{mod}} \in N^*\). \(S\) satisfies (Q2) iff \(\forall i \in A, \forall t \in N, \text{if } \forall j \in Inf(i), d_{\text{mod}}(j) \leq L_{\text{mod}}, \text{then } d_{\text{mod}}(i) \leq L_{\text{mod}}.\)

\(S\) satisfies (Q3) iff \(\forall i, j \in A, \forall t \in N, \text{if } d_{\text{extr}}(i) \leq d_{\text{extr}}(j) \text{ then } \|\text{Mod}(B_i^t)\| \leq \|\text{Mod}(B_j^t)\|\).  

\(S\) satisfies (Q4) iff \(\forall i \in A, \forall t \in N, \text{if } \forall j \in Inf(i), d_{\text{extr}}(j) \leq d_{\text{extr}}(i) \text{ then } d_{\text{extr}}(i) \leq d_{\text{extr}}(i)\).

\(S\) satisfies (Q5) iff \(\forall i \in A \text{ such that its first influencer } j \text{ satisfies } d_{\text{extr}}(j) \geq L_{\text{extr}}, d_{\text{extr}}(j) \leq L_{\text{extr}} \text{ and } d_{\text{extr}}(i) \geq L_{\text{extr}} \text{ then } d_{\text{extr}}(i) \geq L_{\text{extr}}.\)

\(L_{\text{lim}} \in N\). \(S\) satisfies (Q5) iff \(\forall i \in A, \forall t \in N, d_{\text{lim}}(i) = \max_{j \in Inf(i)} d_{\text{extr}}(j) \leq L_{\text{lim}}.\)

\(L_{\text{lim}} \in N\). \(S\) satisfies (Q5) iff \(\forall i \in A, \forall t \in N, \text{if } d_{\text{lim}}(i) = \min_{j \in Inf(i)} d_{\text{extr}}(j) \leq L_{\text{lim}}.\)
4 EXTREME OPINIONS AND EXTREMIST AGENTS

In this section, we consider that extreme opinions are strong opinions i.e. formulas which have “few” models. Moreover, we adopt a binary definition of extremism by considering that people are divided into two distinct groups: the extremist ones and the moderate ones.

4.1 Definitions

Definition 9 (Extreme Opinions). Let $R$ be a given integer closer to 1 than to $2^{|d|}$. An opinion $o$ is extreme iff $1 \leq \text{Mod}(o) \leq R$.

The choice of the threshold $R$ will depend on the application. But $R$ has to be much smaller than the number of interpretations in the language. Moreover, inconsistent opinions are not considered as extreme. For instance, consider that the two letters of the language are $a, b$. If $R = 1$ then $a \land b$, $a \land \neg b$, $\neg a \land b$, $\neg a \land \neg b$ are the extreme opinions.

As for extremist agents, we define them as agents which opinions are extreme. Moreover, an agent is moderate when it is not extremist. This leads to the following definition:

Definition 10 (Extremist, Moderate). At step $s$, agent $i$ is extremist iff $B_i^s$ is an extreme opinion. Otherwise it is moderate.

Example 2. Consider two propositional letters $a, b$ and assume that at a given step $t$ opinions are: $B_i^t = a \lor b$, $B_j^t = a$, $B_k^t = a \land b$. If $R = 1$ then only $k$ is extremist. If $R = 2$ then $j$ and $k$ are extremist.

4.2 Formal Analysis

In the following, we prove some results on extremism diffusion and in particular we focus on properties described in section 3.

The following proposition gives a description of extremist agents from a syntactical point of view.

Proposition 1. Assume that the propositional letters are $a_1, \ldots, a_n$. Consider an IODS $S = (A, \mu, B, \text{Inf})$. Agent $i$ is extremist at step $t$ iff $B_i^t \equiv \bigwedge_{k=1}^{N_i} l_{p,k} \land \ldots \land l_{n,k}$ with $\forall k \in [1, N_i]$, $\forall p \in [1, n_i]$, $l_{p,k}$ a literal corresponding to $a_p$, ($l_{p,k} = a_p$ or $l_{p,k} = \neg a_p$) and $N \leq R$.

Thus, the opinion of an extremist agent is equivalent to a disjunction of less than $R$ conjunctions of all the literals.

The following proposition states that an agent which main influencer is extremist at some step will be extremist at the next step.

Proposition 2. In an IODS $S = (A, \mu, B, \text{Inf})$, for $i \in A$ with $\text{Inf}(i) = \{j < i, \ldots\}$, for $t \in N$, if $j$ is extremist at step $t$, then $i$ is extremist at step $t+1$.

As a consequence, a self-confident agent (see definition 4) which is extremist at some step will remain extremist ever after:

Proposition 3. In an IODS $S = (A, \mu, B, \text{Inf})$, for $i \in A$ with $\text{Inf}(i) = \{i < j, \ldots\}$, if $\exists t \in N$, $i$ is extremist at $t$, then $\forall s \geq t$, $i$ is extremist at $s$.

The following proposition states that an agent which $k$-th influencer has an opinion consistent with the merging of the ones of the previous influencers at step $t$ and which $k$-th influencer is extremist at step $t$ will be extremist at step $t + 1$.

Proposition 4. In an IODS $S = (A, \mu, B, \text{Inf})$, for $i \in A$ with $\text{Inf}(i) = \{j_1 < i, \ldots, j_k < i, \ldots\}$, for $t \in N$, if $\Delta(t)(B_{j_1}^t \land \ldots \land B_{j_k}^t)$ is consistent and $j_k$ is extremist at step $t$, then $i$ is extremist at step $t + 1$.

More generally, an agent will be extremist at step $t + 1$ iff for some $k$, the merging of the $k$ first influencers’ opinions at step $t$ has less than $R$ models.

Proposition 5. In an IODS $S = (A, \mu, B, \text{Inf})$, for $i \in A$ with $\text{Inf}(i) = \{j_1 < i, \ldots, j_k < i, \ldots\}$. Let $t \in N$. If $\exists k \in N$, such that $|\Delta(t)(B_{j_1}^t \land \ldots \land B_{j_k}^t)| \leq R$ then, $i$ is extremist at $t + 1$. Otherwise, it is moderate at step $t + 1$.

The following proposition states that a self-confident extremist agent spreads extremism in its sphere of influence.

Proposition 6. Let $S = (A, \mu, B, \text{Inf})$ an IODS and $i \in A$ extremist at step $t$ with $\text{Inf}(i) = \{i < j, \ldots\}$. $\exists t' \geq t$, $\forall s \geq t'$, $\forall j \in \text{Sphere}(i)$, $j$ is extremist at step $s$.

The last proposition shows which are the properties formalized in section 3 that are satisfied.

Proposition 7. Any IODS satisfies (P1), (P3) and (P4). (P2) and (P5) are not always satisfied.

5 EXTENSION TO DEGREES OF EXTREMISM

In this section, we extend the notion of extreme opinions according to a more relative point of view.

5.1 Definitions

Definition 11 (Degree of Extremism of Opinions). An opinion $\varphi$ is extreme at degree $\frac{2^n - |\text{Mod}(\varphi)|}{2^n - 1}$ and moderate at degree $1 - \frac{2^n - |\text{Mod}(\varphi)|}{2^n - 1}$ with $N$ the number of propositional letters.
Example 3. For instance, if we consider the letters \(a\) and \(b\) then \(a \land b\) is extreme at degree 1 and \(a \lor \neg b\) is extreme at degree \(1/3\).

Definition 12 (Degree of Extremism of Agents). At step \(s\), the degree of extremism of agent \(i\) is the degree at which \(B^s_i\) is extreme. It is denoted \(d^s_{\text{ext}}(i)\). Similarly, at step \(s\), the degree of moderation of agent \(i\) is the degree at which \(B^s_i\) is moderate. It is denoted \(d^s_{\text{mod}}(i)\).

Example 4. An agent which current opinion is \(a \land b\) is more extremist than agent \(j\) iff \(d^s_{\text{ext}}(i) \geq d^s_{\text{ext}}(j)\). At step \(s\), agent \(i\) is more moderate than agent \(j\) iff \(d^s_{\text{mod}}(i) \geq d^s_{\text{mod}}(j)\).

5.2 Formal Analysis

This first proposition states that an agent is more extremist at step \(t\) \((t > 0)\) than its first influencer at step \(t - 1\).

Proposition 8. Consider an IODS \(S = (A, \mu, B, \text{Inf})\), and an agent \(i \in A\) with \(\text{Inf}(i) = \{j_1 \prec_i \ldots \prec_i j_n\}\). Then \(\forall t > 0, d^t_{\text{ext}}(i) \geq d^{t-1}_{\text{ext}}(j)\).

In particular, the degree of extremism of a self-confident agent can only increase.

The following proposition is a generalization of the previous one. It states that an agent is more extremist at some step \(t\) than its \(k\)-th influencer at the previous step if the opinions of this influencer is consistent with the merging of the opinions of the previous ones.

Proposition 9. Consider an IODS \(S = (A, \mu, B, \text{Inf})\) and an agent \(i\) with \(\text{Inf}(i) = \{j_1 \prec_i \ldots \prec_i j_n\}\). Then, \(\forall t > 0, \forall k \in [1, n], \text{if } \Delta_\mu(B_{j_1} \prec \ldots \prec B_{j_n}) \land B^t_{j_k}\) is consistent and \(j_k\) is an extremist agent at step \(t\) then \(d^{t+1}_{\text{ext}}(i) \geq d^t_{\text{ext}}(j_k)\).

The following proposition states that an agent will be extremist at degree at least \(d\) at step \(t + 1\) if one of its influencers \(j\) has an opinion at step \(t\) such that during the opinion updating process, the merging of the previous influencers’ opinions has less than \(2^N - (2^N - 1)d\) models at minimal distance from \(B^t_j\) with \(N\) the number of letters in the language.

Proposition 10. In an IODS \(S = (A, \mu, B, \text{Inf})\) with \(A = \{a_1, \ldots, a_N\}\), for \(i \in A\), \(t \in \mathbb{N}\), with \(\text{Inf}(i) = \{j_1 \prec_i \ldots \prec_i j_n\}\), \(i\) is of degree of extremism at least \(d\) at step \(t + 1\) iff \(\exists k \in [1, n], |\text{Mod}(\Delta_\mu(B_{j_1} \prec \ldots \prec B_{j_n}))| \leq 2^N - (2^N - 1)d\).

The following proposition states that a self-confident extremist agent spreads extremism in its sphere of influence.

Proposition 11. Let \(S = (A, \mu, B, \text{Inf})\) an IODS and \(i \in A\) with \(\text{Inf}(i) = \{i \prec \ldots \prec j_n\}\). \(\forall t \in \mathbb{N}, \exists t' \geq t, \forall s \geq t', \forall j \in \text{Sphere}(t)\), \(d^s_{\text{ext}}(j) \geq d^t_{\text{ext}}(i)\).

The last proposition shows which are the propositions formalized in section 3 that are satisfied.

Proposition 12. Any IODS satisfies (Q1), (Q3), (Q4) and (Q7). (Q2), (Q5) and (Q6) are not always satisfied.

6 EXPERIMENTS

In this section, we focus on simulations realized with Netlogo. First, we address the question of generating IODS corresponding to real social networks. For that, we review some propositions made in graph theory during the last decades. Then, we adapt them to our context by using a notion of distances between opinions.

6.1 Graph Theory Bases

One of the most used models of graph is the one of Erdős-Rényi. It is a model of random graph (see (Easley and Kleinberg, 2010)).

Definition 14 (Erdős-Rényi Graph). Given a number of nodes \(n\) and an integer \(m\). An Erdős-Rényi Graph is any graph obtained by selecting randomly \(m\) edges among the \(2^n\) possible ones.

Another model of graph that is widely used is the model of Watts-Strogatz. This model has been made to describe the phenomenon of Small-World or “six degrees of separation” highlighted by Milgram (Milgram, 1967). This psychologist established through an experiment the theory that a message can be transmitted from one person to one another by passing by an average of six friends. The Small-World theory is commonly formalized (Easley and Kleinberg, 2010; Prettjeohn et al., 2002; Watts and Strogatz, 1998) as follows:

Definition 15 (Small-World). A graph \(G\) is said Small-World if it satisfies:

1. \(G\) is connected.
2. \(G\) is sparse: the average degree of the nodes \(k\) is low compared to the number of nodes \(n\), \(k \ll n\).
3. \(G\) is decentralized: the maximal degree of the nodes \(k_{\text{max}}\) is low compared to the number of nodes \(n\), \(k_{\text{max}} \ll n\).
4. The characteristic path \(L\) (the average number of nodes traversed by a short path between two nodes) is close to the one of a random graph with...
the same number of nodes $n$ and the same average degree $k$, $L \approx L_{\text{random}} \sim \frac{\ln(n)}{\ln(k)}$.

(v) The clustering coefficient $C$ (the probability that two nodes $i$ and $j$ are connected given that they share a common neighbor) is high compared to the one of a random graph with the same number of nodes $n$ and the same average degree $k$, $C \gg 1$.

One can notice that Erdős–Rényi graphs as random graphs have low characteristic paths by definition.

The following model, from (Easley and Kleinberg, 2010) and adapted from a model generally attributed to Watts and Strogatz, define Small-World graphs:

**Definition 16** (Rank-Based Friendship Graph). Given a number of nodes $n$, a threshold $r$, an exponent $q$ and a dimension $d$, the nodes are randomly distributed in a space of dimension $d$. Rank-Based Friendship Graph is obtained by going as follows:

For each node $i$, we rank the other nodes according to their distances to $i$ and we break ties with a chosen method. There will be an edge from a node $j$ to the node $i$ with probability $\frac{1}{Z \sum_{j \in \text{neighbors}(i)} \text{rank}(j)^q}$ being the rank of $j$ in $i$’s neighbors and $Z$ a coefficient of normalization, $Z = \sum_{j=1}^{n} \frac{1}{\text{rank}(j)^q} = \frac{\sum_{j=1}^{n}}{\sum_{j=1}^{n}}$.

**6.2 Distances between Opinions**

Considering a pseudo-distance $d$ between two interpretations, there are many ways to characterize how close two opinions are. (Eiter and Mannila, 1997) highlights several pseudo-distances between sets of points, we adapt some to opinions as follows:

**Definition 17** (Pseudo-Distances on opinions). Let $o_1$ and $o_2$ be two propositional formulas.

- **Sum of minimum distances**: $d_{\text{summin}}(o_1, o_2) = \frac{1}{2} (\sum_{w \in \text{Mod}(o_1)} D(w, o_2) + \sum_{w \in \text{Mod}(o_2)} D(w, o_1))$

- **Hausdorff distance**: $d_{\text{Haus}}(o_1, o_2) = \max(\max_{w \in \text{Mod}(o_1)} D(w, o_2), \max_{w \in \text{Mod}(o_2)} D(w, o_1))$

- **Surjection distance**: $d_{\text{surj}}(o_1, o_2) = \min_{q \in \text{Mod}(o_2) \setminus \text{Mod}(o_1)} \sum_{w_1, w_2 \in q} d(w_1, w_2)$ where the set of surjections from the larger set between $\text{Mod}(o_1)$ and $\text{Mod}(o_2)$ on the other.

- **Link distance**: $d_{\text{link}}(o_1, o_2) = \min_{L \in \text{Link}(o_1, o_2)} \sum_{w_1, w_2 \in R} d(w_1, w_2)$ with $L_{o_1, o_2}$ the set of linkings. A linking $R \subseteq \text{Mod}(o_1) \times \text{Mod}(o_2)$ satisfies $\forall w_1 \in \text{Mod}(o_1), \exists w_2 \in \text{Mod}(o_2), (w_1, w_2) \in R$ and $\forall w_2 \in \text{Mod}(o_2), \exists w_1 \in \text{Mod}(o_1), (w_1, w_2) \in R$.

But, as we consider relations of influence that are not symmetric, dropping the property of symmetry and considering weaker functions than pseudo-distances will also give interesting IODS to study. We thus define Difference functions as follows:

**Definition 18** (Difference of Opinions).

- $d_{\text{max}}(o_1, o_2) = \max_{w_1 \in \text{Mod}(o_1), w_2 \in \text{Mod}(o_2)} d(w_1, w_2)$
- $d_{\text{min}}(o_1, o_2) = \min_{w_1 \in \text{Mod}(o_1), w_2 \in \text{Mod}(o_2)} d(w_1, w_2)$
- $d_{\text{maxmin}}(o_1, o_2) = \max_{w_1 \in \text{Mod}(o_1)} \min_{w_2 \in \text{Mod}(o_2)} d(w_1, w_2)$
- $d_{\text{sum}}(o_1, o_2) = \sum_{w_1 \in \text{Mod}(o_1), w_2 \in \text{Mod}(o_2)} d(w_1, w_2)$

In the following we will indifferently refer to the previously defined distances by lack of a better term.

**6.3 Models**

Here we adapt the previous models of graphs to IODS and explain how we construct them for our simulations. In the following we take an integrity constraint being a tautology.

The first model we adapt is the one of the random graph defined by Erdős and Rényi. The following definition shows how we construct Erdős and Rényi IODS. Notice that we add a parameter, the number of self-confident agents, which is an interesting variable to study.

**Definition 19** (Erdős–Rényi-Based IODS). Given the parameters num-letters, num-nodes, num-links and num-self-confident, the IODS is constructed as follows:

We begin by creating num-nodes agents, each of them has a random opinion in a language of num-letters. Then, we create num-links relations of influence by choosing randomly an influencer and an influenced agent (potentially the same). The influencers are ordered according to the order of creation of the relation of influence, the sooner a relation of influence would have been created the more influencing it is. Finally, each agent with no influencers will become dogmatic and, if necessary, we add relations of self-influence until we have num-self-confident self-confident agents (dogmatic agents included). We pick randomly an agent and if it does not already influence himself we make it self-confident by putting it as its main influencer (the order of the other influencers remains unchanged).

This second model adapts the model of Rank-Based Friendship by considering a distance between opinions instead of a physical distance as for the graph model.
**Definition 20** (Rank-Based Influenceship IODS). Given the parameters num-letters, num-nodes, opinions-distance, q and num-self-confident, the IODS is constructed as follows:

We begin by creating num-nodes agents, each of them has a random opinion in a language of num-letters. Then, we fill a matrix with the distances between every couple of agents according to the distance between their opinions and computed with the distance opinions-distance. For each agent i we have a list l_i of all the agents (i included) sorted according to their distances to i. If two agents j_1 and j_2 are at the same distance of i, then the tie will be randomly solved. Each agent j will be an influencer of i with probability \( \frac{1}{Z \cdot \text{rank}(j)^q} \), rank(i) being the rank of i in l_i and Z being a coefficient of normalization, \( Z = \sum_{l=1}^{\text{num-nodes}} \frac{1}{\text{rank}(j)^q} \). The influencers of i are ordered as in l_i. Finally, if necessary, we add relations of self-influence such as we have num-self-confident self-confidence agents (dogmatic agents included). We pick randomly an agent and if it is not already self-confident we make it so by putting it as its main influencer (the order of the other influencers remains unchanged).

The third model is a variant of the previous one, here an agent will be influenced by the m agents that have the closest opinions from its own one for a given integer m.

**Definition 21** (Deterministic Rank-Based Influenceship IODS). Given the parameters num-letters, num-nodes, opinions-distance, m and num-self-confident, the IODS is constructed as follows:

We begin by creating num-nodes agents, each of them has a random opinion in a language of num-letters. Then, we fill a matrix with the distances between every couple of agents according to the distance between their opinions and computed with the distance opinions-distance. For each agent i, we conserve the m closest agents to i to be its influencers. If two agents j_1 and j_2 are at the same distance of i, then the tie will be randomly solved. The influencers of i are ordered according to their distances to i. Finally, we add relations of self-influence such as we have num-self-confident self-confidence agents (dogmatic agents included). We pick randomly an agent and if it is not already self-confident we make it so by putting it as its main influencer (the order of the other influencers remains unchanged).

The fourth model is a generalization of the Rank-Based Influenceship in which we have in addition to the distance between opinions a physical distance along a circle. The influencers of an agent i are ordered according to the distance between their opinions and the one of i.

**Definition 22** (Opinions and Physical Rank-Based Influenceship IODS). Given the parameters num-letters, num-nodes, opinions-distance, r, q and num-self-confident, the IODS is constructed as follows:

We begin by creating num-nodes agents, each of them has a random opinion in a language of num-letters. We fill a matrix with the distances between every couple of agents according to the distance between their opinions and computed with the distance opinions-distance. For each agent i we have a list l_i of all the agents (i included) sorted according to their distances to i. If two agents j_1 and j_2 are at the same distance of i, then the tie will be randomly solved. Each agent j will be an influencer of i with probability \( \frac{1}{Z \cdot \text{rank}(j)^q} \), rank(i) being the rank of j in l_i and Z a coefficient of normalization \( Z = \sum_{l=1}^{\text{num-nodes}} \frac{1}{\text{rank}(j)^q} \). At the previous influencers we add influencers that are physically close. Indeed, all the agents will be placed on a circle. The agents that are separated on the circle from an agent i by less than r agents will influence i. The influencers of i are ordered as in l_i (according to the distance between opinions). Finally, if necessary, we add relations of self-influence such as we have num-self-confident self-confidence agents (dogmatic agents included). We pick randomly an agent and if it is not already self-confident we make it so by putting it as its main influencer (the order of the other influencers remains unchanged).

To study and compare the results between the different models and distances, we carried out several simulations with the same settings. Furthermore, in order to do comparable and reproducible experiments we chose some values of seeds for the random operations in Netlogo. Seeds allow to have the same results in the same order for random operations when we repeat the simulations. The values we study are the number of extremist agents for \( R = 1 \), the average number of models per agents which is proportional to the average degree of moderation and the number of dogmatic agents.

In the simulations we present here, we have taken the following values: seed \( [0, 100, 200] \), num-letters \( \{3, 4, 5, 6\} \), num-nodes \( \{10, 60, 110, 160, 210\} \), num-self-confident \( \{0, 50, 100, 150, 200, 210\} \). For the three models using ranks we tested \( d_{\text{summin}}, d_{\text{Haus}}, d_{\max}, d_{\min}, d_{\text{maximum}}, d_{\text{sum}} \) with the drastic pseudo-distance between interpretations. For the Erdős-Rényi-Based model we took num-links varying from 10 to 2000 with an increment of 50. For the Rank-
Based Influenceship and the Opinions and Physical Rank-Based Influenceship models we took the values \{1, 2, 3, 4, 5\} for \(q\), for the Deterministic Rank-Based Influenceship model we took \(m\) in \{1, 2, 3, 4\} and for the Opinions and Physical Rank-Based Influenceship model we took \(r\) in \{1, 2, 3, 4, 5\}. The pseudo-distance used to compute the Importance-based Merging operator is the drastic one.

First of all, with the drastic pseudo-distance, we can notice that the different distances we used have particular behaviors. \(d_{\text{summin}}(\alpha_1, \alpha_2) \in [0, 2^{\text{num-letters}}]\) and \(d_{\text{summin}}\) favors relations of influence between agents which opinions have models very close according to \(D\) in average. \(d_{\text{sum}}(\alpha_1, \alpha_2) = 0\) iff \(\alpha_1 \equiv \alpha_2\). \(d_{\text{summin}}\) favors relations of influence between agents which opinions have no models very far from one another. So, in the case of the drastic pseudo-distance, it favors relations between agents that have the same opinion. So, for the number of letters and agents we will consider, as such a case is unlikely the relations of influence will be mostly random. \(d_{\text{max}}(\alpha_1, \alpha_2) = 0\) iff \(\exists \omega \in \text{Mod}(\alpha_1) \cap \text{Mod}(\alpha_2)\). So, as having two agents with only one model and the same model is very unlikely for the number of letters and agents we consider, \(d_{\text{max}}\) favors random relations of influence and it will be interesting to compare the results obtained with this distance and the ones obtained with the other distances. \(d_{\text{min}}(\alpha_1, \alpha_2) = 0\) iff \(\exists \omega \in \text{Mod}(\alpha_1) \cap \text{Mod}(\alpha_2)\). Then, for a given agent \(i\), \(d_{\text{min}}\) favors relations of influence that are from agents that share a model with \(i\’s\) opinion but that are otherwise random. \(d_{\text{maxmin}}(\alpha_1, \alpha_2) = 0\) iff \(\exists \omega \in \text{Mod}(\alpha_1) \cap \text{Mod}(\alpha_2)\). So, \(d_{\text{maxmin}}\) favors relations of influence from an agent \(i\) to an agent \(j\) such that all the models of \(B_j\) are models of \(B_i\). \(d_{\text{sum}}(\alpha_1, \alpha_2) \in [0, 2^{\text{num-letters}}]\) and \(d_{\text{sum}}\) favors relations of influence from agents which opinions have the less models.

For the Erdös-Rényi-Based model (see figure 1), we have several peaks of the average number of models and of the number of dogmatic agents, corresponding to having low \(\text{num-links}\). Indeed, in these cases, there are potentially more agents that are not influenced by other agents and that keep their initial opinions. Furthermore, we can see that the dogmatic agents are almost the only agents that are not extremist and thus contribute the more to the average number of models. So, the diffusion of extremism depends a lot on the ratio between \(\text{num-links}\) and \(\text{num-nodes}\), the more there are relations of influence the more the agents will become extremist. We can only notice that the peaks of average number of models are higher and higher according to the increasing of the number of letters. Another experiment in which we took 200 agents and much more relations of influence (up to 7000) showed that for more than 2000 there are very few simulations with non-extremist agents.

For the Rank-Based Influenceship model (see figure 2), we have several plateaus higher and higher according to the increasing of the number of letters. Furthermore, there are cases with very low numbers of extremist agents and without very much dogmatic agents. Then, we can notice that there are big differences according to the distance we use. Indeed, a thorougher analysis highlights that the biggest peaks are with \(d_{\text{summin}}\) and then with \(d_{\text{Haus}}\) and \(d_{\text{maxmin}}\). \(d_{\text{max}}\) and \(d_{\text{min}}\) cause some lesser peaks when \(q\) gets bigger (more than 3) and \(d_{\text{sum}}\) causes very small peaks for \(q = 5\). For \(q = 1\) almost all the agents are extremist whatever the distance we use. It can be explained by the fact that the lesser \(q\) is the more likely relations of influence are to be created, furthermore for \(q\) high enough the distance used matter less then even \(d_{\text{sum}}\).
that in the other cases spread extremism may be used to create an IODS where they may remain some moderate agents. But, according to (Easley and Kleinberg, 2010) in the case of graphs, the Rank-Based Friendship generates graphs the closest of reality for \( q = 1 \). Furthermore, when the number of agents increases, the average number of models decreases because more relations of influence may be created. One can notice that with this model \( d_{sumin} \) and \( d_{Haus} \) particularly favor moderation. So, having influencers with opinions for which each model is close of one of us model or for which each model is not far of any of our model favor moderation. But, we can notice that with \( d_{Haus} \) agents are much less dogmatic than with \( d_{sumin} \).

For the Deterministic Rank-Based Influenceship model (see figure 3), extremism spreads more and more when \( m \) gets bigger. Moreover, this time there is much more differences according to the distance we used because the ranking is more important in the choice of the influencers than before. Then, only \( d_{sumin} \) keeps many non-extremist agents when \( m \) is at its highest. Indeed, this distance characterizes the best the similarity between opinions, the first agents in the ranking of an agent \( i \) actually have opinions that share many models with the one of \( i \) and it often is \( i \) itself. Thus, when \( m = 1 \), we have almost only dogmatic agents with \( d_{sumin} \). When \( m \) gets higher than 3 only models with \( d_{sumin} \) keep moderate agents.

For the Opinions and Physical Rank-Based Influenceship model (see figure 4), we have very few non-extremist agents even with \( d_{sumin} \) and even less when \( r \) increases. It is due to the fact that here there cannot be any dogmatic agent (contrary to the case of the Rank-Based Influenceship) and that an agent may have influencers with very different opinions (contrary to the case of the Deterministic Rank-Based Influenceship).

In all the simulations, the number of letters does not affect the proportion of extremist agents. The number of nodes affects the proportion of extremist agents for the Erdős-Rényi-Based model because of our definition of the model, in fact it is the ratio between the number of agents and the number of relations of influence that truly matters. It also has an influence for Rank-Based Influenceship model and the Opinions and Physical Rank-Based Influenceship model because it increases the average number of in-
fluencers.

For summarizing, we have that for the models of generation of IODS, the ones which spread extremism the less are the Rank-Based Influenceship when $q$ is very high and the Erdős-Rényi-Based when num-links is much lower than num-nodes. But, those models have many dogmatic agents, on the other hand, the Deterministic Rank-Based Influenceship spreads extremism very little with $d_{\text{summin}}$ and a small $m$ and without many dogmatic agents. For the distances, it is $d_{\text{summin}}$ that spreads extremism the less because it favors relations of influence from agents with opinions sharing many models and it spreads extremism less than $d_{\text{max}}$ (the random one). At the opposite, $d_{\text{sum}}$ spreads extremism very well by creating hubs, agents with very few models that influence a lot of agents. $d_{\text{min}}$ spreads extremism a little less because it is less random, there is a constraint on one model. So, with the Importance-Based Merging Operator, the extremism spreads very well when the most extremist agents are very influential and much less when agents are influenced by agents with opinions similar to its own in the sense of they share many models. So, what makes that an agent remains moderate is the fact that he is influenced by agents which opinions share many models between them and that he does not have too many influencers. Having many self-confident agents favor extremism spreading with the Erdős-Rényi-Based model as it increases the average number of influencers but in the other models it favors moderation. Indeed, in this case the agents keep opinions close to their initial ones and so agents’ influencers keep close opinions.

We can notice that, in every simulation, we reached the convergence very quickly in general in less than 5 updates.

It would have been interesting to test the models for much larger numbers of agents to increase the probabilities we have deemed negligible in our study of the distances for instance. Indeed, the Small-World phenomenon is considered interesting for very large number of nodes i.e. billions of nodes (see (Watts, 1999)) but the computation time that would be needed only for models of thousands of agents is very important.

Furthermore, other simulations with Hamming pseudo distance both for the computation of the distances between opinions and the update of the opinions gave similar results. Notwithstanding, extremism spreads slightly much, in average 0.8 less models per agents and 9% less extremist agents. This can mainly be explain by the fact that $\text{Min}_{\mathcal{d}_H, \varphi_1 \prec \varphi_2, \text{Mod}(\mu)}$ contains generally less models than $\text{Min}_{\mathcal{d}_H, \varphi_1 \prec \varphi_2, \text{Mod}(\mu)}$ as the second one keeps all the models of $\varphi_1$ if $\varphi_1$ and $\varphi_2$ are inconsistent. The only type of IODS that spreads less extremism in this case is the Rank-Based Influenceship model, in average there are 2 more models per agents and 10% less extremist agents. But, it can be explained by the fact that there are twice more (15% more) dogmatic agents, the hamming pseudo-distance allows a more accurate ranking of the agents and thus, it is less likely that agents with very different opinions influence an agent. It appears that this accuracy is all the more significant that the number of letters is important. However, the first agents in the rankings do not change a lot, so the Deterministic Rank-Based Influenceship model spreads more extremism. Another noticeable difference are for $d_{\text{Haus}}$ and $d_{\text{summin}}$ which spread extremism much less in the three Rank-Based Influenceship models.

What we can notice is that the more relations of influence there are, the more extremism spreads. And, the more influencers of agents have close opinions, the less extremism spreads. This result can be interpreted as follows: When someone makes its own mind by taking into account the opinions of many people it considers as reliable or experts on the matter and with different opinions then, it will be very sure
of its new opinion as it is a compromised between the opinions of many experts. And so, this person will become extremist according to our definition.

7 CONCLUDING REMARKS

This paper focused on modelling extreme opinion diffusion when opinions are modelled as propositional formulas. It can be extended according to several directions.

First, we could add a dynamic aspect in the different types of IODS, by changing the relations of influence through time as it is often done in the usual models (Crawford et al., 2013; Christoff and Hansen, 2015; Chau et al., 2014; Deffuant et al., 2002). It will be especially interesting with the rank-based models where the ranking of the influencers is based on the distances between opinions. Since opinions change through time these distances also change and computing new rankings could be done.

Another aspect that may be complexified is the definition of extreme opinions. In the definitions we considered here, the main parameter is the number of the models of an opinion. They do not take into account what opinions are about. For instance, according to this model, the opinion using pesticides is safe is as extreme as growing tomatoes is easy. Taking the domain into account would allow us to distinguish some sensitive letters and to use them for a more refined definition of extremism. In the agriculture domain, having a strong position towards pesticides (pro or cons) is obviously more noticeable than having a strong position towards tomatoes.

Similarly, we could consider a more complex definition of extremism which would define extremist agents as agents which opinions are close to some referential extreme opinions. In the agriculture example, an agent which thinks that with caution, using pesticides is safe is more extremist than an agent which thinks with water, growing tomatoes is easy because its opinion is not far from the sensitive opinion using pesticides is safe.

ACKNOWLEDGEMENTS

We thank the anonymous reviewers whose comments helped us to improve the paper.

REFERENCES


