Reliability-Based Preference Dynamics: Lexicographic Upgrade *

Fernando R. Velázquez-Quesada[1]

[1] *Institute for Logic, Language and Computation, Universiteit van Amsterdam.*
F.R.VelazquezQuesada@uva.nl

Abstract

This paper models collective decision-making scenarios by using a priority-based aggregation procedure, the so-called *lexicographic* method, in order to represent a form of reliability-based 'deliberation'. More precisely, it considers agents with a preference ordering over a set of objects and a reliability ordering over the agents themselves, providing a logical framework describing the way in which the public and simultaneous announcement of the individual preferences leads to individual preference upgrade. The main results are the definitions of this *lexicographic upgrade* for diverse types of reliability relations (in particular, the preorder and total preorder cases), a sound and complete axiom system for a language describing the effects of such upgrades, and the definitions for non-public variations.

Keywords: multiagent systems; preference; reliability; preference change; lexicographic; modal logic; dynamic epistemic logic; aggregation; deliberation.

1 Introduction

Suppose four friends want to watch a movie, with the options being 1, 2 and 3. As it frequently happens in similar situations, the preferences of the friends are different. So, what can they do in order to fix a ranking of the three options, and thus decide which movie they will see? More generally, what can a set of agents do in similar collective decision making scenarios?

The friends might try to put together their individual preferences. This can be done by looking for an *aggregation* procedure: a process through which the individual preferences are combined into a single one. As stated in Endriss [42], *"when a group needs to make a decision, we are faced with the problem of aggregating the views of the individual members of that group into a single collective view that adequately reflects the 'will of the people' "*. Finding appropriate aggregation procedures is the fundamental aim of research fields as social choice theory [4; 5; 6] as well as preference/belief change/merge/aggregation (e.g., Konieczny

*`lexj-2017-04-28(ferrvq).tex`. To appear in the *Journal of Logic and Computation*.

and Pino Pérez [75, 76]; Grüne-Yanoff and Hansson [64]; Gabbay et al. [49]; Liu [88]; Konieczny and Pino Pérez [77]). Examples of such procedures are the different voting systems used across different countries.[1]

Another alternative the friends have is to allow each one of them to argue for her preferences, trying in this way to change the opinions of the others. This is called *deliberation*, a process through which the individuals share publicly not only their own preferences but also the justifications they have for them, aiming at influencing one another's opinions. Several authors have pointed to the merits of the deliberative process, which makes people reflect on their preferences and thus influences possible changes [41; 67; 65; 23; 37; 30; 66]. Some authors have even argued that with such a public debate *"there would not be any need for an aggregation mechanism, since a rational decision would tend to produce unanimous preferences"* [41]. The campaign process that takes place in most countries before an election might be seen as some form of deliberation, as it allows some form of public debate through which ideas and reasons for supporting the candidates are exposed.[2]

In some sense, the deliberative process is the ideal one: agents share not only their own preferences but also the justifications they have for them, thus putting all the available information 'on the table', and then proceed to a thoughtful weighing of options based on the use of logic and reason. Even if the repetition of such sharing and discussing process does not lead to full unanimity (an obvious possibility), it has been argued [94; 93; 38; 56] that the deliberative process can lead to preference profiles whose properties allows/facilitates the use of aggregation procedures.

However, in real-life scenarios, different circumstances are combined to guarantee that ideal deliberative processes occur rarely. One of these circumstances is the fact that, even though agents might get to know one another's preferences, certain limitations prevent them from knowing the reasons/justifications for them. Another is that typically, and for diverse reasons, different agents might influence a specific individual in different ways.

Because of these and other 'real-life' constrains, attempts to carry out a deliberative process might actually end up in situations in which an individual's thoughts, opinions, feelings, and behaviours are influenced by (the actual, imagined or implied presence of) others. Examples of such phenomena include not only the famous *informational cascades* [11; 21; 104; 40] (cf. *the bandwagon effect*; Leibenstein [83]; Nadeau et al. [95]; Altman [1]), but also *peer pressure* [39], *pluralistic ignorance* [73; 78], *false consensus* [97] and others studied in economics, computer science and social and behavioural sciences. All these different ways in which the preferences/beliefs/behaviour of a group of agents influence the preferences/beliefs/behaviour of an individual might not be completely 'rational', but nevertheless describe the way 'real' agents behave in 'real-life' situations.

This paper provides a formal logical framework for studying a process that

[1] It should be noted, however, that the aim of a typical voting system is to collectively select one option, and not to create a full collective preference.

[2] Still, such process is not deliberation in the proper sense: before the voting, typically not all the voters will have the opportunity to announce their preferences and justifications in public forums, and even if they do, not all of these announcements will be heard by everyone.

deals with collective decision making scenarios following a 'deliberative' approach, but incorporating the two limitations described above: a knowledge of the preferences but not of its reasons, and some agents having more influence than others on each individual.[3] The reason for incorporating these particular limitations is that they reflect very common real-life circumstances. For the first, indeed in most cases we only get to know one another's actions/opinions/preferences/beliefs without knowing the reasons behind them, typically because of restrictions in the 'communication channel' (i.e., lack of time and/or space). A famous consequence of this is that, paraphrasing a popular quote, even though we tend to judge ourselves by our intentions, we tend to judge others only by their actions. For the second, the fact that some agents might have more influence than others is also a common feature that ideally can be seen as agents acknowledging the different expertise individuals might have, but which also allows to describe 'irrational' *herd behaviour* situations in which emotions ("I like a better than b, so I prefer what a chooses over what b chooses") come into play.[4]

More precisely, the goal of the present proposal is the formal study of the way individual agents might change their preferences based not only on the preferences of one another but also on the reliability ordering each one of them assigns to the group. Technically, the models used here will then consist, besides a domain W and a set of agents $\mathtt{A}$, of both a *preference* relation $\leq_i \subseteq (W \times W)$ for each agent $i \in \mathtt{A}$ and a *reliability* relation $\preccurlyeq_i \subseteq (\mathtt{A} \times \mathtt{A})$ for each such i. While $\leq_i$ represents the preference ordering i has over the available options W, $\preccurlyeq_i$ represents the reliability ordering i has over the involved agents $\mathtt{A}$.[5] In this way, the procedure through which the announcement of the individual preferences leads each agent i to change her preferences can be seen as a function f whose parameters are not only the (current/announced) preferences $(\leq_1, \ldots, \leq_{|\mathtt{A}|})$ but also her own reliability relation $(\preccurlyeq_i)$. In this sense this proposal can be understood as a qualitative version of DeGroot [31], where the author presents a quantitative model describing how a group might reach a consensus and form a common subjective probability distribution by revealing their individual distributions to one another and pooling their opinions. In such model, each individual i has not only a subjective probability distribution (corresponding to this work's preference relation) but also a weight she assigns to the distribution of each individual when she carries out her revision (corresponding to this work's reliability relation).

As the reader has surely noticed, the definition of the mentioned f will play the most important role through this work. Again, it represents intuitively the way an agent i will adjust her own preferences according to all the individual preferences and the reliability she assigns to them. Again, technically it can be defined as a function which takes as parameters $|\mathtt{A}|$ preference orderings over W

[3] There might be additional limitations: for example, the communication might not be public, or the preferences might not be fully announced. Some of these situations can be seen as particular cases of the framework that will be introduced, and they are briefly explored in Section 4.

[4] Some people have pointed out that, because of this difference in influence, not all agents are equal, contrary to what should happen in ideal deliberative procedures. However, the degree of influence of each agent is not global but rather individual: the most influential agent for a given individual might not be the most influential agent for a different one. Thus, there is no 'most important agent(s)'; only 'most important agent(s)' *from a given agent's point of view*.

[5] Subsection 2.1 discusses other options for representing preference and, in particular, reliability.

and a reliability ordering over $\mathtt{A}$ (i.e., $|\mathtt{A}|$ orderings over W and a priority ordering over them), returning then a preference ordering over W. This function can be defined in many different ways: for example, i could adopt directly the preference ordering of her most reliable agent, or she might keep her old ordering, using that of her most reliable agent to 'break ties' in equally-preferable zones. She might even put her most reliable agent's most preferred worlds above the rest, using her old ordering to break ties in both zones.

In cases as the present one, a typical choice is to use the *lexicographic method*. This method consists, roughly speaking, in using the order with the highest priority first, then using the order with the second highest priority to 'break ties' in equally-preferred zones, then using the order with the third highest priority to 'break ties' in the equally-preferred zones that still remain, and so on; it is similar to the method used for ordering words in dictionaries. The lexicographic method has been used not only in social choice theory [46; 47; 22; 80; 81] but also in artificial intelligence (preferential logics: Ryan [100]; Grosof [61]; Schobbens [102]; Ryan [101]; defaults in the setting of circumscription: Lifschitz [85, 86]; beliefs: Brandenburger et al. [24]; Dekel et al. [32]; belief revision: Ryan [101]), and it is the most natural one for defining a function that *aggregates* a collection of orders based on a priority order over them. Indeed, the first two examples mentioned before (adopt the preference ordering of the most reliable agent, keep the old ordering using that of the most reliable agent to 'break ties') are particular instances of the lexicographic method in which the priority ordering only considers a subset of agents [53]. The third example is not lexicographic in the strict sense, as it cannot be defined as a lexicographic aggregation of the original preferences [52], but it still has a lexicographic flavour: it is the lexicographic aggregation of, first, an 'arbitrary' preference relation (the one placing her most reliable agent's most preferred worlds above the rest) and then the agent's original preference relation.

Summarizing, this proposal studies non-completely rational but nevertheless realistic situations in which, after the public announcement of all the agents' individual preferences, each individual agent might change her own preferences based not only on what has been announced but also on the reliability order she assigns to the members of the group. All agents will change her individual preferences using the described lexicographic method, but even though all of them will use the just announced preferences as input, each one of them will use *her own* reliability ordering for prioritizing the preferences to be aggregated, thus potentially producing different results. This approach for studying collective decision making scenarios borrows some ideas from the deliberative tradition, as it acknowledges that the preferences of individuals can be influenced by the preferences of the rest (albeit with the mentioned limitations). Still, ideas from the aggregation literature are also clearly present, the main one being the use of the lexicographic method for defining each individual's preferences after each announcement.

The work is structured as follows. Section 2 presents formally the basic model sketched above for representing both the individual's preferences about objects and the relative reliability they assign to one another. It also recalls a formal language for describing these structures as well as an axiom system characterising its validities. Then, after presenting a definition of the lexicographic method for the cases in which the reliability ordering is a total order (a straightforward variation of the one presented in Ghosh and Velázquez-Quesada [53]) and a partial order [2], Section 3 introduces its preorder and total preorder versions. The section also presents a modality for describing the model after all agents have upgraded their preferences according to the method, together with recursion axioms for its axiomatisation. Section 4 explores 'non public' variations of the lexicographic upgrade operation, and Section 5 connects this proposal with related works. Finally, Section 6 closes by summarizing the work and discussing possible ways to extend it.

Within the text, proofs of propositions and theorems are sometimes only sketched; in such cases, the formal proof can be found in the Appendix.

On preference change Despite the fact that diverse preference influence phenomena have been studied extensively in the literature, not only in social and behavioural sciences, but also in economics and in computer science (see Section 5), there might be still a conviction among some people that human preferences ultimately do not change. In this respect, this paper's best strategy is to refer the interested reader to the discussion in Grüne-Yanoff and Hansson [63]. But even if the reader remains skeptic about the preference-change idea, this paper's proposal might still be meaningful, as agent a's revised 'preferences' can be also understood not as a's 'new' preferences but rather simply as some preferences she decides to announce in order to keep the 'discussion' ongoing. Accordingly, the reliability ordering, singling out the agents whose preferences a trusts the most, can be rather understood as a 'pleaseability' ordering, singling out the agents a wants to *please* the most.

2 Basic definitions

Throughout this paper, let $\mathtt{A}$ be a *finite non-empty* set of agents with $|\mathtt{A}| = n$. The following is the basic definition of this framework: the structure representing each agent's relative preferences and reliability over, respectively, worlds (i.e., the available options) and one another.

Definition 2.1 (*PR* frame) A *preference and reliability (PR) frame* F is a tuple $\langle W, \{\leq_i, \preccurlyeq_i\}_{i \in \mathtt{A}} \rangle$ where

- W is a finite non-empty set of worlds;
- $\leq_i \subseteq (W \times W)$ is a preorder (a reflexive and transitive relation), agent i's *preference relation* over worlds in W ($u \leq_i v$ is read as " for agent i, world v is at least as preferable as world u");
- $\preccurlyeq_i \subseteq (\mathtt{A} \times \mathtt{A})$ is a preorder, agent i's *reliability relation* over agents in $\mathtt{A}$ ($j \preccurlyeq_i j'$ is read as " for agent i, agent j' is at least as reliable as agent j"). ◄

In the context of aggregation and deliberation it is common to assume at least two different agents; this proposal follows such assumption.

Here are further useful definitions. First, given any domain D, any element $d \in D$ and any binary relation $R \subseteq (D \times D)$, the set $[R\rangle(d)$ contains those elements in D that are R-reachable from d, that is,

$$[R\rangle(d) := \{e \in D \mid Rde\}$$

Then,

Definition 2.2 Let $\leq \subseteq (W \times W)$ be a preference relation (preorder) over W.

- The relation $< \subseteq (W \times W)$, with $u < v$ read as "u is less preferred than v" ("v is more preferred than u" or, simply, "v is preferred over u"), is defined as $u < v$ iff$_{def}$ $u \leq v$ and $v \nleq u$.
- The relation $\sim \subseteq (W \times W)$, with $u \sim v$ read as "u and v are comparable", is defined as $u \sim v$ iff$_{def}$ $u \leq v$ or $v \leq u$.
- The relation $\simeq \subseteq (W \times W)$, with $u \simeq v$ read as "u and v are equally preferred", is defined as $u \simeq v$ iff$_{def}$ $u \leq v$ and $v \leq u$.

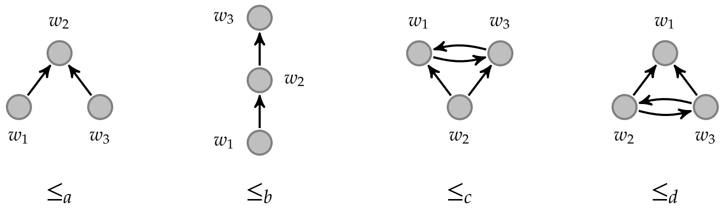
Then, one more definition. For any set $U \subseteq W$, the binary relation $\mathrm{Id}_U \subseteq (W \times W)$ is the identity relation on U, that is, $\mathrm{Id}_U := \{(u, u) \in (W \times W) \mid u \in U\}$.◀

Definition 2.3 Let $\preccurlyeq \subseteq (\mathtt{A} \times \mathtt{A})$ be a reliability relation (preorder) over $\mathtt{A}$.

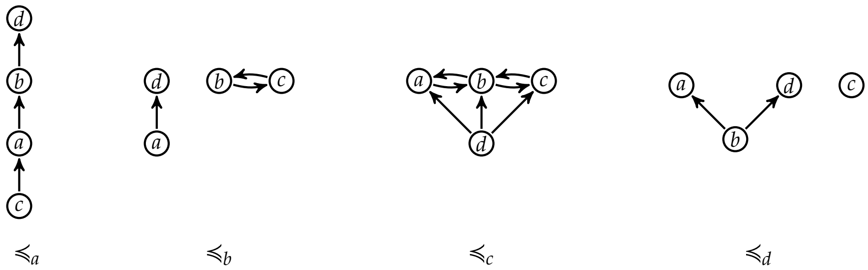
- The relation $\prec \subseteq (\mathtt{A} \times \mathtt{A})$, with $j \prec k$ read as "j is less reliable than k" (or as "k is more reliable than j"), is defined as $j \prec k$ iff$_{def}$ $j \preccurlyeq k$ and $k \not\preccurlyeq j$.
- The relation $\approx \subseteq (\mathtt{A} \times \mathtt{A})$, with $j \approx k$ read as "j and k are comparable", is defined as $j \approx k$ iff$_{def}$ $j \preccurlyeq k$ or $k \preccurlyeq j$.
- The relation $\cong \subseteq (\mathtt{A} \times \mathtt{A})$, with $j \cong k$ read as "j and k are equally reliable", is defined as $j \cong k$ iff$_{def}$ $j \preccurlyeq k$ and $k \preccurlyeq j$.
- The set $\mathrm{Mnl}_{\preccurlyeq} \subseteq \mathtt{A}$, containing those elements in $\mathtt{A}$ that are $\preccurlyeq$-minimal, is defined as $\mathrm{Mnl}_{\preccurlyeq} := \{k \in \mathtt{A} \mid \text{there is no } j \in \mathtt{A} \text{ such that } j \prec k\}$. As it is well-known, when $\preccurlyeq$ is also total[6], $\mathrm{Mnl}_{\preccurlyeq}$ actually contains those elements in $\mathtt{A}$ that are $\preccurlyeq$-minimum, i.e., those that are $\preccurlyeq$-below all elements in $\mathtt{A}$ (formally, $\mathrm{Mnl}_{\preccurlyeq} = \{k \in \mathtt{A} \mid \text{for all } j \in \mathtt{A},\ k \preccurlyeq j\}$); in such case, $\mathrm{Mnl}_{\preccurlyeq}$ will be denoted by $\mathrm{Mnm}_{\preccurlyeq}$. If $\preccurlyeq$ is additionally antisymmetric[7], then $\mathrm{Mnm}_{\preccurlyeq}$ becomes a singleton whose single element will be denoted by $\mathrm{mn}_{\preccurlyeq}$.
- The set $\mathrm{Mxl}_{\preccurlyeq} \subseteq \mathtt{A}$, containing those elements in $\mathtt{A}$ that are $\preccurlyeq$-maximal, is defined as $\mathrm{Mxl}_{\preccurlyeq} := \{k \in \mathtt{A} \mid \text{there is no } j \in \mathtt{A} \text{ such that } k \prec j\}$. Similar to the previous case, when $\preccurlyeq$ is also total, $\mathrm{Mxl}_{\preccurlyeq}$ actually contains those elements in $\mathtt{A}$ that are $\preccurlyeq$-maximum, i.e., those that are $\preccurlyeq$-above all elements in $\mathtt{A}$ (formally, $\mathrm{Mxl}_{\preccurlyeq} = \{k \in \mathtt{A} \mid \text{for all } j \in \mathtt{A},\ j \preccurlyeq k\}$); in such case, $\mathrm{Mxl}_{\preccurlyeq}$ will be denoted by $\mathrm{Mxm}_{\preccurlyeq}$. If $\preccurlyeq$ is additionally antisymmetric, then $\mathrm{Mxm}_{\preccurlyeq}$ becomes a singleton, whose single element will be denoted by $\mathrm{mx}_{\preccurlyeq}$.
- A set $\mathtt{B} \subseteq \mathtt{A}$ is an $\preccurlyeq$-*cluster* when $\preccurlyeq$ is a maximal universal relation over $\mathtt{B}$, that is, when
  (i) $i \preccurlyeq j$ for all $i, j \in \mathtt{B}$ (or, in other words, the restriction of $\preccurlyeq$ to $\mathtt{B}$ is a universal relation, $\preccurlyeq \cap (\mathtt{B} \times \mathtt{B}) = \mathtt{B} \times \mathtt{B}$), and
  (ii) for all $i \in \mathtt{A} \setminus \mathtt{B}$, the set $\mathtt{B} \cup \{i\}$ does not satisfy the previous point (or, in other words, $\mathtt{B} \subset \mathtt{B}'$ implies $R \cap (\mathtt{B}' \times \mathtt{B}') \neq \mathtt{B}' \times \mathtt{B}'$). ◀

[6]For every $j, k \in \mathtt{A}$, $j \preccurlyeq k$ or $k \preccurlyeq j$ (or both).
[7]For every $j, k \in \mathtt{A}$, $j \preccurlyeq k$ and $k \preccurlyeq j$ imply $j = k$.

Example 2.1 The situation described in this work's opening paragraph can be represented by a *PR* frame in which the set of agents is $\mathtt{A} = \{a, b, c, d\}$ and the set of worlds is $W = \{w_1, w_2, w_3\}$ (with each world representing one movie). Suppose the individual preferences are as described in the following diagrams, with reflexive and transitive edges omitted.



Observe how, for example, while w_1 and w_3 are not comparable for a ($w_1 \nsim_a w_3$), the latter is more preferred than the former for b ($w_1 <_b w_3$), they are equally preferred for c ($w_1 \simeq_c w_3$) and the former is more preferred than the later for d ($w_3 <_d w_1$). Moreover, suppose reliability is given as in the following diagrams (again, omitting reflexive and transitive edges).



Now observe how, for example, while c is less reliable than a for a herself ($c \prec_a a$), they are incomparable for b and d ($c \not\approx_b a$ and $c \not\approx_d a$) but equally reliable for c ($c \approxeq_c a$). Note also how $\mathrm{Mxl}_{\preccurlyeq_a} = \mathrm{Mxm}_{\preccurlyeq_a} = \{d\}$, as $\preccurlyeq_a$ is total. This is not the case for $\preccurlyeq_b$, and thus $\mathrm{Mxl}_{\preccurlyeq_b} \neq \mathrm{Mxm}_{\preccurlyeq_b}$, as $\mathrm{Mxl}_{\preccurlyeq_b} = \{b, c, d\}$ but $\mathrm{Mxm}_{\preccurlyeq_b} = \varnothing$. Moreover, while $\{b, c\}$ is not a $\preccurlyeq_a$-cluster (its elements are not pairwise $\preccurlyeq_a$-equally reliable), it is a $\preccurlyeq_b$-cluster; nevertheless, it is neither a $\preccurlyeq_c$-cluster (its elements are pairwise $\preccurlyeq_c$-equally reliable, but so are those of its strict super set $\{a, b, c\}$) nor a $\preccurlyeq_d$-cluster (same reason as for $\preccurlyeq_a$). Finally, $[\preccurlyeq_a\rangle(b) = \{b, d\}$, $[\preccurlyeq_b\rangle(b) = \{b, c\}$, while $[\prec_c\rangle(b) = \varnothing$ and $[\prec_d\rangle(b) = \{a, d\}$. ◀

2.1 On preference and reliability

As mentioned earlier, this work models situations in which, after individual preferences are announced, each agent changes her preferences according to what has been announced and the relative reliability she assigns to all the agents (including herself: she might consider herself as more reliable than some agents but also as less reliable than some others).

In the defined frames, the agents' preferences are represented by a binary relation, a strategy used by several works in the formal study of preferences (see, e.g., von Wright [105]; Arrow et al. [5, 6]; Grüne-Yanoff and Hansson [64]; Liu [88] and further references therein). Such relation is typically assumed to be at least reflexive and transitive, as in this work. Thus, when presented with options u and v, each agent i can give one of four answers: ***(i)*** she prefers v over u, $u <_i v$, ***(ii)*** she prefers u over v, $v <_i u$, ***(iii)*** she considers both equally preferable, $u \simeq_i v$, or ***(iv)*** she considers them incomparable, $u \nsim_i v$ (i.e., she abstain from any judgement). Note how, since incomparability is an option, maximum worlds (those that are at least as preferred as any other element) might not exist; nevertheless, given the finiteness of the domain, there are always maximals ones (those that are not less preferred than some other element). An analogous statement is true for minimums and minimals.

The concept of *reliability* between agents requires a deeper discussion. It is related to the notion of *trust*, which has been important within artificial societies (e.g., Falcone et al. [44, 45]; Golbeck [55]), and for which there are several proposals for its formal representation. It is worthwhile to discuss, albeit briefly, how *reliability* as discussed here relates to *trust*.

Though there are proposals that define trust as an attitude of an agent who believes that another agent has a given property [27; 33; 34; 43], a common understanding of this concept is as "agent i trusts agent j's judgement about φ" (called "trust on credibility" in Demolombe [33]). And, while there are approaches that define trust in terms of other attitudes, as knowledge, beliefs, intentions and goals (e.g., Demolombe [33]; Herzig et al. [70]), others define it as a semantic primitive, typically by means of a *neighbourhood function* $N_{i,j} : W \to \wp(\wp(W))$ that assigns, to every pair of agents i, j in every world w, a set of sets of worlds $N_{i,j}(w)$. In such frameworks, it is said that agent i trusts agent j's judgement about φ at world w if and only if the set of worlds in W where φ holds is in $N_{i,j}(w)$ [84].[8] The notion of trust is not represented with normal modal semantics in order to avoid closure under logical consequence: that agent i trusts agent j's judgement about some formula does not imply that i also trusts j on the formula's logical consequences.

In contrast, reliability as discussed here is closer to the notion of trust of Holliday [71], where it is understood as an ordering among sets of sources of information (cf. the discussions in Cantwell [26]; Goldman [59]). One noticeable difference between the more standard representation of trust (on credibility) and what is called reliability here is then that the former parametrises trust with a formula (which can be understood as a topic or area of expertise). Nevertheless, the key distinction is that the latter does not yield *absolute* judgements ("i relies on j's judgement [about φ]"), but only *comparative* ones ("for i, agent k is at least as reliable as agent j"). For the purposes of this work, such comparative judgements will suffice, as the main goal is to describe the way the relative reliability/trust in a collection of sources affects the way the information the sources provide is assimilated.[9]

In this proposal, the reliability relation is asked to be a reflexive and transitive relation. As mentioned, these are the two natural requirements for an ordering, but when the ordering is intended to represent some form of priority, two other properties are frequently chosen: totality and antisymmetry. Asking

[8] Some variants deal with *graded trust*, as Demolombe and Liau [35]; Lorini and Demolombe [90]; Lorini et al. [91].

[9] Note how it is also possible to define reliability relative to a particular topic by requiring a relation $\preccurlyeq_i^U$ for each agent i and each set of worlds $U \subseteq W$. Then, $j \preccurlyeq_i^U k$ can be understood as "when discussing subject φ, agent k is at least as reliable as agent j for agent i", with φ a formula that is true exactly in the worlds in U. This paper works with 'plain' reliability in order to not deviate the attention from its main topic.

for these two extra properties would amount to force every agent to select, among any pair of agents, a single most reliable one. This paper's choice is, instead, a more realistic setting that allows for any two agents not only to be incomparable (as the relation does not need to be total) but also to be equally reliable (as the relation does not need to be antisymmetric).

Finally, note how the approaches mentioned above as well as the present one consider reliability/trust as a concept that is part of the system's definition. This is because, again, the aim is to understand how the announcement of the agents' preferences and the relative reliability/trust they assign to one another affects individual preferences. There are, of course, other possibilities. For example, one could understand reliability/trust as a concept that arises *from the system's behaviour*, using then the way the information is assimilated in order to define the absolute/comparative trust in the sources. Such approach would understand this notion not statically, but rather dynamically.

2.2 A formal language

Throughout this paper, let $\mathtt{P}$ be a countable set of atomic propositions.

Definition 2.4 (Language $\mathcal{L}$) Formulas φ, ψ ($\mathcal{L}^f$) and relational expressions π, σ ($\mathcal{L}^r$) of the language $\mathcal{L}$ are given by

$$\varphi, \psi ::= \top \mid p \mid j \sqsubseteq_i k \mid \neg\varphi \mid \varphi \vee \psi \mid \langle\pi\rangle\varphi$$
$$\pi, \sigma ::= 1 \mid \leq_i \mid \geq_i \mid ?(\varphi, \psi) \mid -\pi \mid \pi \cup \sigma \mid \pi \cap \sigma$$

with $p \in \mathtt{P}$ and $i, j, k \in \mathtt{A}$. Other propositional constants ($\bot$), other Boolean connectives ($\wedge, \rightarrow, \leftrightarrow$) and the dual modal universal operators $[\pi]$ are defined as usual ($[\pi]\varphi := \neg\langle\pi\rangle\neg\varphi$ for the latter). Define also, for any relational expression π, the modal operator $\boxed{\pi}$ as $\boxed{\pi}\,\varphi := [-\pi]\,\neg\varphi$. All these abbreviations will simplify both the writing of formulas and the presentation of the axiom system. ◂

The set of formulas of $\mathcal{L}$ contains the always true formula ($\top$), atomic propositions (p) and formulas for describing the agents' reliability relations ($j \sqsubseteq_i k$), and it is closed under negation ($\neg$), disjunction ($\vee$) and modal operators of the form $\langle\pi\rangle$ with π a relational expression. The set of relational expressions contains the constant 1 (the global relation, as it will follow from its semantic definition), the preference relations ($\leq_i$) and their respective converse ($\geq_i$; Prior [96]; Burgess [25]; Goldblatt [57]) and an additional construction of the form $?(\varphi, \psi)$ with φ and ψ formulas of the language [52], and it is closed under Boolean operations over relations (the so-called *Boolean modal logic*; Gargov and Passy [50]; Lutz and Sattler [92]).

The following two definitions establish what a model is and how formulas of $\mathcal{L}$ are interpreted over them.

Definition 2.5 (*PR* model) A *PR model* M is a tuple $\langle F, V\rangle$ where F is a *PR* frame and $V : \mathtt{P} \rightarrow \wp(W)$ is a *valuation function*. The domain of a model M will be denoted sometimes as $\mathcal{D}_M$. ◂

Definition 2.6 (Semantic interpretation) For the semantic interpretation, let $M = \langle W, \{\leq_i, \preccurlyeq_i\}_{i\in\mathtt{A}}, V\rangle$ be a *PR* model. The function $[\![\cdot]\!]^M : \mathcal{L}^f \rightarrow \wp(W)$, from formulas in $\mathcal{L}$ to subsets of W, and the function $\langle\!\langle\cdot\rangle\!\rangle^M : \mathcal{L}^r \to \wp(W \times W)$, from relational expressions in $\mathcal{L}$ to binary relations over W, are defined inductively and simultaneously in the following way.

$$
\begin{aligned}
[\![\top]\!]^M &:= W & [\![j \sqsubseteq_i k]\!]^M &:= \begin{cases} W & \text{if } j \preccurlyeq_i k \\ \varnothing & \text{otherwise} \end{cases}\\
[\![p]\!]^M &:= V(p) \\
[\![\neg\varphi]\!]^M &:= W \setminus [\![\varphi]\!]^M \\
[\![\varphi \vee \psi]\!]^M &:= [\![\varphi]\!]^M \cup [\![\psi]\!]^M \\
[\![\langle\pi\rangle\,\varphi]\!]^M &:= \left\{w \in W \mid \langle\!\langle\pi\rangle\!\rangle^M_w \cap [\![\varphi]\!]^M \neq \varnothing\right\}
\end{aligned}
$$

with $\langle\!\langle\pi\rangle\!\rangle^M_w$ the set of worlds reachable from w via $\langle\!\langle\pi\rangle\!\rangle^M$,[10] and

$$
\begin{aligned}
\langle\!\langle 1\rangle\!\rangle^M &:= W \times W & \langle\!\langle -\pi\rangle\!\rangle^M &:= (W \times W) \setminus \langle\!\langle\pi\rangle\!\rangle^M \\
\langle\!\langle \leq_i\rangle\!\rangle^M &:= \leq_i & \langle\!\langle \pi \cup \sigma\rangle\!\rangle^M &:= \langle\!\langle\pi\rangle\!\rangle^M \cup \langle\!\langle\sigma\rangle\!\rangle^M \\
\langle\!\langle \geq_i\rangle\!\rangle^M &:= \{(v,u) \in (W \times W) \mid u \leq_i v\} & \langle\!\langle \pi \cap \sigma\rangle\!\rangle^M &:= \langle\!\langle\pi\rangle\!\rangle^M \cap \langle\!\langle\sigma\rangle\!\rangle^M \\
\langle\!\langle ?(\varphi,\psi)\rangle\!\rangle^M &:= [\![\varphi]\!]^M \times [\![\psi]\!]^M
\end{aligned}
$$

Note, in particular, how $\langle\!\langle ?(\varphi,\psi)\rangle\!\rangle^M$ is the set of those pairs $(u,v) \in (W \times W)$ such that u satisfies φ and v satisfies ψ.[11] As usual, a formula φ is *true* at world w in model M when $w \in [\![\varphi]\!]^M$, and it is true at M when $[\![\varphi]\!]^M = \mathcal{D}_M$. As usual, φ is *valid* when $[\![\varphi]\!]^M = \mathcal{D}_M$ for every model M. ◀

As a consequence of the previous definition,

$$
[\![\langle 1\rangle\,\varphi]\!]^M = \begin{cases} W & \text{if } [\![\varphi]\!]^M \neq \varnothing \\ \varnothing & \text{otherwise} \end{cases}
$$

and hence $\langle 1\rangle$ is the global existential modality. Moreover,

$$
[\![[\pi]\,\varphi]\!]^M = \left\{w \in W \,\middle|\, \langle\!\langle\pi\rangle\!\rangle^M_w \subseteq [\![\varphi]\!]^M\right\}, \qquad [\![\boxed{\pi}\,\varphi]\!]^M = \left\{w \in W \,\middle|\, [\![\varphi]\!]^M \subseteq \langle\!\langle\pi\rangle\!\rangle^M_w\right\}
$$

Thus, while $[\pi]$ is a standard universal modality for the relation $\langle\!\langle\pi\rangle\!\rangle^M$, the modality $\boxed{\pi}$ is the window operator [58; 20; 51].

On the language's operators First, note how reliability formulas $j \sqsubseteq_i k$ can be combined with Boolean operators to express more complex reliability relations:

- $(j \equiv_i k) := (j \sqsubseteq_i k) \wedge (k \sqsubseteq_i j)$
- $(j \not\equiv_i k) := \neg(j \equiv_i k)$
- $(j \sqsubset_i k) := (j \sqsubseteq_i k) \wedge \neg(k \sqsubseteq_i j)$
- $(j \not\sqsubseteq_i k) := \neg(j \sqsubseteq_i k) \wedge \neg(k \sqsubseteq_i j)$

Similarly, with relational expressions it is possible to define modalities for relations that can be defined from $\leq_i$ (see Definition 2.2). For example,

[10] Formally, $\langle\!\langle\pi\rangle\!\rangle^M_w := [\![\langle\!\langle\pi\rangle\!\rangle^M\rangle(w)$.

[11] Compare $\langle\!\langle ?(\varphi,\psi)\rangle\!\rangle^M$ with the relation $\langle\!\langle ?\varphi\rangle\!\rangle^M := \{(u,u) \in (W \times W) \mid u \in [\![\varphi]\!]^M\}$ for the traditional *PDL* test operation $?\varphi$ [69].

- $\not\geq_i := - \geq_i$
- $\not\leq_i := - \leq_i$
- $<_i := \leq_i \cap \not\geq_i$
- $>_i := \not\leq_i \cap \geq_i$
- $\simeq_i := \leq_i \cap \geq_i$

and hence

- $\langle \not\geq_i \rangle \varphi := \langle - \geq_i \rangle \varphi$
- $\langle \not\leq_i \rangle \varphi := \langle - \leq_i \rangle \varphi$
- $\langle <_i \rangle \varphi := \langle \leq_i \cap \not\geq_i \rangle \varphi$
- $\langle >_i \rangle \varphi := \langle \not\leq_i \cap \geq_i \rangle \varphi$
- $\langle \simeq_i \rangle \varphi := \langle \leq_i \cap \geq_i \rangle \varphi$

Moreover, the operator $?(\varphi, \psi)$ allows the construction of relations connecting worlds that satisfy the specified formulas. All these relations, modalities and formulas will be useful not only for describing the agents' preferences and reliabilities, but also for simplifying formulas and also for providing axioms for the lexicographic upgrade operation to be introduced in Subsection 3.3.

Example 2.2 Consider the frame described in Example 2.1; let $\mathrm{P} = \{p_1, p_2, p_3\}$ be the set of atomic propositions, and take a valuation V such that

$$V(p_1) := \{w_1\}, \qquad V(p_2) := \{w_2\}, \qquad V(p_3) := \{w_3\}$$

(i.e., each p_i represents the movie i). The language $\mathcal{L}$ can be used to describe the situation. For example, the following formulas are true at the model, as each one of them is true in all the model's worlds:

- $p_2 \rightarrow [<_a] \perp$: world w_2 is one of a's maximals.
- $[1] \langle \leq_a \rangle p_2$: world w_2 is one of a's maximums.
- $p_3 \leftrightarrow [<_b] \perp$: world w_3 is b's unique maximal (and, as $\leq_b$ is total and antisymmetric, it is also her unique maximum).
- $p_2 \rightarrow [>_c] \perp$: world w_2 is one of c's minimals.
- $((p_2 \vee p_3) \leftrightarrow [>_d] \perp) \wedge (p_2 \rightarrow \langle \simeq_d \rangle p_3)$: worlds w_2 and w_3 are d's minimals, and she considers them are equally preferred.
- $(c \sqsubset_a a) \wedge (c \not\sqsubseteq_b a) \wedge (c \equiv_c a) \wedge (c \not\sqsubseteq_d a)$: each agent's opinion about the relative reliability of a and c. ◀

Digression: from preference over objects to preference over properties A *PR* frame represents the agents' preferences over worlds. Such preference over objects can be lifted to an ordering over sets of such objects (i.e., an ordering over object's *properties*) in different ways, as it has been discussed by economic theorists (the standard reference is Barberà et al. [14]) and by logicians [18; 54; 19; 88]. For example, one can say that the set of objects satisfying the property ψ (the set of ψ-objects) is at least as preferable as the set of objects satisfying the property φ (the set of φ-objects) when *there is* a ψ-object that the original ordering considers at least as preferable as *some* φ-object (a $\exists\exists$ preference of ψ over φ; see below). But one can be more drastic and say that the set of ψ-objects is at least as preferable as the set of φ-ones when the original ordering considers *every* ψ-object at least as preferable as *every* φ-one (a $\forall\forall$ preference of ψ over φ). In general, the quantification combination gives raise to the following possibilities:

$\varphi \leq^{\exists\exists} \psi$ iff$_{def}$ *there is* a φ-object w and *there is* a ψ-object u such that $w \leq u$
$\varphi \leq^{\forall\exists} \psi$ iff$_{def}$ *for every* φ-object w *there is* a ψ-object u such that $w \leq u$
$\varphi \leq^{\exists\forall} \psi$ iff$_{def}$ *there is a* φ-object w such that $w \leq u$ *for every* ψ-object u
$\varphi \leq^{\forall\forall} \psi$ iff$_{def}$ $w \leq u$ *for every* φ-object w and *every* ψ-object u

which, by using a first-order language and indexing the original preference relation over objects, can be also defined as

$$\begin{array}{lll}
\varphi \leq_i^{\exists\exists} \psi & \text{iff}_{def} & \exists w.\big(\varphi(w) \wedge \exists u.(\psi(u) \wedge w \leq_i u)\big) \\
\varphi \leq_i^{\forall\exists} \psi & \text{iff}_{def} & \forall w.\big(\varphi(w) \rightarrow \exists u.(\psi(u) \wedge w \leq_i u)\big) \\
\varphi \leq_i^{\exists\forall} \psi & \text{iff}_{def} & \exists w.\big(\varphi(w) \wedge \forall u.(\psi(u) \rightarrow w \leq_i u)\big) \\
\varphi \leq_i^{\forall\forall} \psi & \text{iff}_{def} & \forall w.\big(\varphi(w) \rightarrow \forall u.(\psi(u) \rightarrow w \leq_i u)\big)
\end{array}$$

The language $\mathcal{L}$ is expressive enough to define formulas depicting such orderings. The first two cases, taken from the aforementioned references, are straightforward: besides the global modalities, they only require modalities for the agent's preference relation.

Proposition 2.1 *Let $M = \langle W, \{\leq_i, \preccurlyeq_i\}_{i \in \mathsf{A}}, V \rangle$ be a PR model. Then,*

- $[\![\langle 1 \rangle(\varphi \wedge \langle \leq_i \rangle \psi)]\!]^M = W$ *iff* $\varphi \leq_i^{\exists\exists} \psi$
- $[\![[1](\varphi \rightarrow \langle \leq_i \rangle \psi)]\!]^M = W$ *iff* $\varphi \leq_i^{\forall\exists} \psi$

Proof. See the Appendix (page 35). ■

Thanks to the defined relational operators (in particular, thanks to the existence of modalities for $<_i$ and $>_i$), the remaining cases can be also expressed within $\mathcal{L}$ with a caveat: both $\varphi \leq_i^{\exists\forall} \psi$ and $\varphi \leq_i^{\forall\forall} \psi$ imply that the original preference over objects is *total*, as the first requires for *every* ψ-object to be $\leq_i$-reachable from the given φ-one, and the second requires for *every* ψ-object to be $\leq_i$-reachable from every φ-one. Thus, one possibility is simply to assume that indeed the given *PR* model is based on a frame in which every preference relation is total.

Proposition 2.2 *Let $M = \langle W, \{\leq_i, \preccurlyeq_i\}_{i \in \mathsf{A}}, V \rangle$ be a PR model in which each $\leq_i$ is a total preorder. Then,*

- $[\![\langle 1 \rangle(\varphi \wedge [>_i] \neg\psi)]\!]^M = W$ *iff* $\varphi \leq_i^{\exists\forall} \psi$
- $[\![[1](\psi \rightarrow [<_i] \neg\varphi)]\!]^M = W$ *iff* $\varphi \leq_i^{\forall\forall} \psi$

Proof. See the Appendix (page 35). ■

Another alternative is to modify the definitions of $\leq_i^{\exists\forall}$ and $\leq_i^{\forall\forall}$ to make them suitable for the more general case of non-total preference relations.

Proposition 2.3 *Let $M = \langle W, \{\leq_i, \preccurlyeq_i\}_{i \in \mathsf{A}}, V \rangle$ be a PR model. Define*

$$\begin{array}{lll}
\varphi \leq_i^{\exists\forall'} \psi & \textit{iff}_{def} & \exists w.\big(\varphi(w) \wedge \forall u.((\psi(u) \wedge w \sim_i u) \rightarrow w \leq_i u)\big) \\
\varphi \leq_i^{\forall\forall'} \psi & \textit{iff}_{def} & \forall w.\big(\varphi(w) \rightarrow \forall u.((\psi(u) \wedge w \sim_i u) \rightarrow w \leq_i u)\big)
\end{array}$$

Then,

- $[\![\langle 1\rangle(\varphi \wedge [>_i]\neg\psi)]\!]^M = W$ *iff* $\varphi \leq_i^{\exists\forall'} \psi$
- $[\![[1](\psi \rightarrow [<_i]\neg\varphi)]\!]^M = W$ *iff* $\varphi \leq_i^{\forall\forall'} \psi$

Proof. See the Appendix (page 36). ∎

Axiom system Here is an axiom system characterising formulas of $\mathcal{L}$ valid in *PR* models [52].

Theorem 1 *Table 1 provides a sound and complete axiom system (with i any agent and π any relational expression) for $\mathcal{L}$ with respect to PR models.*

Proof. The first five blocks of the table are known to be sound and complete for the fragment of the language they take care of: ***(i)*** axioms and rules in the first block take care of propositional validities; ***(ii)*** those in the second establish that every modality is normal; ***(iii)*** axioms in the third state that $\leq_i$ is a reflexive and transitive relation; ***(iv)*** those in the fourth establish that $\geq_i$ is the converse of $\leq_i$ [96; 25]; ***(v)*** axioms and rules of the fifth characterise validities involving Boolean relational operations [50], with (?) the axiom for the relational test operator [52]. Showing that axioms in the sixth block characterise reflexivity and transitivity for the relations $\preccurlyeq_i$ is straightforward. ∎

$\vdash \varphi$ for every propositional tautology φ		From $\vdash \varphi$ and $\vdash \varphi \rightarrow \psi$ infer $\vdash \psi$	
$\vdash [\pi](\varphi \rightarrow \psi) \rightarrow ([\pi]\varphi \rightarrow [\pi]\psi)$	*(K_π)*	From $\vdash \varphi$ infer $\vdash [\pi]\varphi$	*(N_π)*
$\vdash \varphi \rightarrow \langle\leq_i\rangle\varphi$	*($T_\leq$)*	$\vdash \langle\leq_i\rangle\langle\leq_i\rangle\varphi \rightarrow \langle\leq_i\rangle\varphi$	*($4_\leq$)*
$\vdash \varphi \rightarrow [\leq_i]\langle\geq_i\rangle\varphi$	*($Con1_\leq$)*	$\vdash \varphi \rightarrow [\geq_i]\langle\leq_i\rangle\varphi$	*($Con2_\leq$)*
$\vdash \varphi \rightarrow \langle 1\rangle\varphi$	*(T_E)*	$\vdash \langle 1\rangle\langle 1\rangle\varphi \rightarrow \langle 1\rangle\varphi$	*(4_E)*
$\vdash \varphi \rightarrow [1]\langle 1\rangle\varphi$	*(5_E)*		
$\vdash [1]\varphi \leftrightarrow ([\pi]\varphi \wedge \boxed{\pi}\,\neg\varphi)$	*(1_1)*	$\vdash \boxed{1}\,\top$	*(1_2)*
$\vdash \langle ?(\psi_1, \psi_2)\rangle\varphi \leftrightarrow (\psi_1 \wedge \langle 1\rangle(\psi_2 \wedge \varphi))$			*(?)*
$\vdash [-\pi]\varphi \leftrightarrow \boxed{\pi}\,\neg\varphi$	*($-_1$)*	$\vdash [\pi]\neg\varphi \leftrightarrow \boxed{-\pi}\,\varphi$	*($-_2$)*
$\vdash \langle\pi \cup \sigma\rangle\varphi \leftrightarrow (\langle\pi\rangle\varphi \vee \langle\sigma\rangle\varphi)$	($\cup$)	$\vdash \boxed{\pi \cap \sigma}\,\varphi \leftrightarrow \left(\boxed{\pi}\,\varphi \wedge \boxed{\sigma}\,\varphi\right)$	($\cap$)
From $\vdash [\pi]\varphi \rightarrow ([\sigma]\varphi \rightarrow [\rho]\varphi)$ infer		$\vdash [\pi]\varphi \rightarrow \left(\boxed{\rho}\,\neg\varphi \rightarrow \boxed{\sigma}\,\neg\varphi\right)$	*(BR)*
$\vdash j \sqsubseteq_i j$ for every agent j			*(reflexivity)*
$\vdash (j \sqsubseteq_i k \wedge k \sqsubseteq_i \ell) \rightarrow j \sqsubseteq_i \ell$ for all agents j, k, ℓ			*(transitivity)*

Table 1: Axiom system for $\mathcal{L}$. w.r.t. *PR* models.

3 Preference dynamics via lexicographic upgrade

In collective decision scenarios, a public announcement of the agents' individual preferences might induce each agent to change her own preferences according to what has been announced and the reliability ordering she assigns to the set of agents. As mentioned in the Introduction, this section studies a particular method for performing this change: the *lexicographic* method. The study is carried out in different phases, according to the properties of the reliability ordering that defines the priority of the preferences to be aggregated.

3.1 Reliability as a total order

Suppose the reliability relation, the one defining the priority of the preference orderings to be aggregated, is a total order (a total, reflexive, transitive and antisymmetric relation). In such cases, any preference orderings $\leq'$ and $\leq''$ can be compared, and one of them will be strictly more reliable than the other. In such cases, the result of the preferences' lexicographic aggregation can be defined in the following way.

Definition 3.1 (Lexicographic upgrade, total order version) Let $\{\leq_i\}_{i \in \mathtt{A}}$ be a finite collection of binary preference orderings over a domain W; let $\preccurlyeq \subseteq (\mathtt{A} \times \mathtt{A})$ be a total order over $\mathtt{A}$ and recall that, in such cases, $\preccurlyeq$ has a unique minimum $\mathrm{mn}_{\preccurlyeq}$. The preference ordering $\leq_{\preccurlyeq} \subseteq (W \times W)$ is defined as

$$u \leq_{\preccurlyeq} v \ \text{iff}_{def} \ \underbrace{\Big(u \leq_{\mathrm{mn}_{\preccurlyeq}} v \ \wedge \bigwedge_{i \in \mathtt{A} \setminus \{\mathrm{mn}_{\preccurlyeq}\}} u \simeq_i v\Big)}_{1} \ \vee \ \underbrace{\bigvee_{j \in \mathtt{A} \setminus \{\mathrm{mn}_{\preccurlyeq}\}} \Big(u <_j v \ \wedge \bigwedge_{k \in [\prec\rangle(j)} u \simeq_k v\Big)}_{2}$$

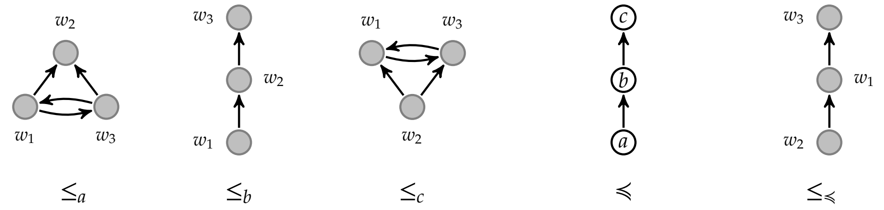
Thus, $u \leq_{\preccurlyeq} v$ holds if this agrees with the least prioritised ordering ($\leq_{\mathrm{mn}_{\preccurlyeq}}$) and for the rest of them u and v are equally preferred (part 1), or if there is an ordering $\leq_j$ with a strict preference for v over u and all orderings with strictly higher priority (those with indexes in $[\prec\rangle(j)$) see u and v as equally preferred (part 2). ◀

As the reliability ordering is a total order, the effect of a lexicographic upgrade can be summarised in a single sentence: the preference ordering with the highest priority is taken as the starting point, and the rest of the orderings are used, hierarchically, to 'break ties' between equally preferred worlds.

In Ghosh and Velázquez-Quesada [53] the authors define a similar operation, called general lexicographic upgrade (*glu*), and used for situations in which the reliability relation is a total order. In a *glu*, both the preference orderings to be aggregated and the priority ordering with which the aggregation will take place are given by what is called a *lexicographic list*. Using such list allows the *glu* to represent other natural lexicographic upgrades, as those that arise when the preferences of some agents in $\mathtt{A}$ are not used for building up the new preference ordering.[12] The just defined lexicographic upgrade is in fact an instance of this *glu* in which the lexicographic list is given directly by the agent's reliability relation.

[12]One example is the mentioned situation in which the agent adopts directly the preference ordering of her most reliable agent. For other examples, see Definitions 7-10 of Ghosh and Velázquez-Quesada [53]. See also Subsection 4.1 for such variations within this framework.

Example 3.1 Consider the preference relations $\leq_a$, $\leq_b$ and $\leq_c$ from Example 2.1 ($\leq_a$ slightly modified), shown below on the left, and consider the application of the lexicographic upgrade of Definition 3.1 following the reliability ordering $\preccurlyeq$ shown below on the center.



The resulting preference ordering, $\leq_\preccurlyeq$, appears above on the right. Following the discussion, it can be seen as the result of refining the preference ordering with the highest priority, $\leq_c$, by using the preference ordering with the second highest priority, $\leq_b$. Since the use of the two topmost preference orderings already makes the resulting relation $\leq_\preccurlyeq$ antisymmetric, the preference ordering with the lowest priority, $\leq_a$, is never consulted. ◂

3.2 Reliability as a partial order: lexicographic rule

Suppose now that the reliability relation is a partial order (a reflexive, transitive and antisymmetric relation). In such cases, two given preference orderings might not be comparable, but if they are, one of them will be strictly more reliable than the other. In such cases, one can rely on the *lexicographic rule* of Andréka et al. [2]. Using this paper's notation, the rule is defined as follows.

Definition 3.2 (Lexicographic rule (cf. Andréka et al. [2])) Let $\{\leq_i\}_{i\in\mathtt{A}}$ be a finite collection of binary preference orderings over a domain W; let $\preccurlyeq$ be a partial order over $\mathtt{A}$.[13] The preference ordering $\leq^{lr}_{\preccurlyeq} \subseteq (W \times W)$ is defined as

$$u \leq^{lr}_{\preccurlyeq} v \qquad \text{iff}_{def} \qquad \bigwedge_{i\in\mathtt{A}} \Big(\underbrace{u \leq_i v}_{1} \vee \underbrace{\bigvee_{j\in[\prec\rangle(i)} u <_j v}_{2} \Big)$$

Thus, $u \leq^{lr}_{\preccurlyeq} v$ holds if and only if every agent in $\mathtt{A}$ agrees with this, or else for every agent who does not agree there is an agent with both a higher priority and a strict preference for v over u. ◂

This lexicographic rule, which requires for $\preccurlyeq$ to be a partial order, looks different from the lexicographic upgrade of Definition 3.1 for cases in which $\preccurlyeq$ is a total order. Nevertheless, as Proposition 3.1 below shows, the two are equivalent when the priority ordering is a *finite total* order.

Proposition 3.1 *Let $\{\leq_i\}_{i\in\mathtt{A}}$ be a finite collection of binary preference orderings over a domain W; let $\preccurlyeq$ be a total order over $\mathtt{A}$. Then, for all $u, v \in W$,*

$$u \leq_{\preccurlyeq} v \qquad \textit{iff} \qquad u \leq^{lr}_{\preccurlyeq} v$$

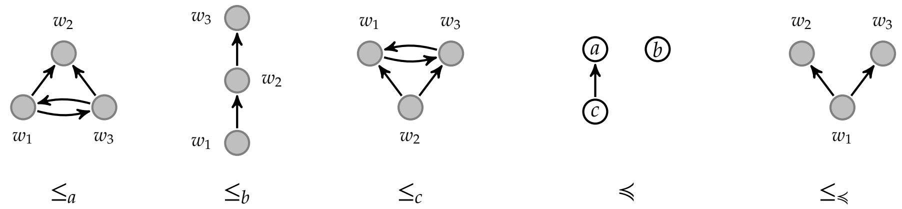
[13] While in Andréka et al. [2] this definition (Definition 2.3 on page 18) allows the collection of preference orderings to be infinite, here the finite set $\mathtt{A}$ is used. Also, in Andréka et al. [2] the orderings to be aggregated are not required to satisfy any property.

Proof. Here is a sketch; for the formal proof, see the Appendix (page 36).

From left to right, $u \leq_{\preccurlyeq} v$ is given by a disjunction (Definition 3.1). If the first disjunct holds, all agents in $\mathtt{A}$ consider v at least as preferable as u, and hence every element of $\mathtt{A}$ satisfies part 1 of Definition 3.2. If the second disjunct holds, there is $j \in \mathtt{A}$ with a strict preference for v over u and with all orderings with higher priority seeing u and v as equally preferred. Hence, while those elements of $\mathtt{A}$ lying $\prec$-below j satisfy part 2, j and those $\prec$-above it satisfy part 1. Since $\preccurlyeq$ is a total order, the previous sentence deals with all elements of $\mathtt{A}$, and hence $u \leq^{lr}_{\preccurlyeq} v$.

From right to left, assume $u \leq^{lr}_{\preccurlyeq} v$. According to Definition 3.2, every agent either considers v at least as preferable as u, or else can $\prec$-see a more reliable agent with a strict preference of v over u. Suppose every agent considers v at least as preferable as u; if, additionally, every agent in $\mathtt{A} \setminus \{\mathrm{mn}_{\preccurlyeq}\}$ considers u at least as preferable as v then, by part 1 of Definition 3.1, $u \leq_{\preccurlyeq} v$. If the additional assumption fails, then there is at least an agent in $\mathtt{A} \setminus \{\mathrm{mn}_{\preccurlyeq}\}$ who does not consider u at least as plausible as v. Given $\mathtt{A}$'s finiteness and $\preccurlyeq$'s properties, among such agents there is a single one with the highest priority, and she satisfies the requirements in part 2 of Definition 3.1; hence, $u \leq_{\preccurlyeq} v$. Suppose now otherwise, i.e., suppose $u \leq_i v$ does not hold for all $i \in \mathtt{A}$; then there is at least one element in $\mathtt{A}$ which does not consider v as at least as plausible as u. Let j be, among such agents, the one with the highest priority, so agents $\prec$-above j consider v at least as plausible as u. From Definition 3.2, there should be at least one agent with higher priority that prefers v strictly over u; let j' be, among such agents, the one with the highest priority, so agents $\prec$-above j' do not prefer v strictly over u. But such agents are also $\prec$-above j; then, they consider v and u equally preferable, and therefore j' satisfies the requirements of part 2 of Definition 3.1; hence, $u \leq_{\preccurlyeq} v$. ■

Example 3.2 Consider the preference relations $\leq_a$, $\leq_b$ and $\leq_c$ from Example 3.1, shown below on the left. Consider now the application of the lexicographic upgrade of Definition 3.2 with the reliability ordering $\preccurlyeq$ shown below on the center.



The resulting preference ordering, $\leq_{\preccurlyeq}$, appears above on the right. Note how there is an edge from w_1 to w_3 because all preference orderings agree with this (i.e., $w_1 \leq_a w_3$, $w_1 \leq_b w_3$, $w_1 \leq_c w_3$). On the other hand, there is an edge from w_1 to w_2 because ***(i)*** a agrees, ***(ii)*** b agrees and ***(iii)*** while c does not agree, there is someone with a strictly higher priority than that of c (namely, a) who places w_1 strictly below w_2. There are no further edges (besides the reflexive ones) because either a or b does not agree with them, and none of them has someone with a strictly higher priority overruling her. ◀

3.3 Reliability as a preorder

In Definition 3.2, the priority ordering is assumed to be a partial order, i.e., a reflexive, transitive and antisymmetric relation. Thus, even though this allows for two different agents to be incomparable, it does not allow for them to have equal priority. Since the reliability relations in a *PR* frame allow equal reliability (as antisymmetry is not required), the main focus of the present work will be the following generalisation.

Definition 3.3 (Lexicographic upgrade, preorder version) Let $\{\leq_i\}_{i\in\mathtt{A}}$ be a finite collection of binary preference orderings over a domain W; let $\preccurlyeq$ be a preorder over $\mathtt{A}$. The relation $\leq_{\preccurlyeq} \subseteq (W \times W)$ is defined as

$$u \leq_{\preccurlyeq} v \ \text{iff}_{def} \ \bigwedge_{i\in\mathtt{A}} \Big(\underbrace{u \leq_i v}_{1} \vee \underbrace{\bigvee_{\mathtt{B}\subseteq[\prec\rangle(i)} \Big(\overbrace{\bigwedge_{j_1,j_2\in\mathtt{B}} j_1 \simeq j_2}^{2.1} \wedge \overbrace{\bigwedge_{k\in\mathtt{A}\setminus\mathtt{B}} \bigvee_{j\in\mathtt{B}} j \not\simeq k}^{2.2} \wedge \overbrace{\bigwedge_{j\in\mathtt{B}} u <_j v}^{2.3} \Big)}_{2} \Big)$$

According to this definition, $u \leq_{\preccurlyeq} v$ holds if and only if every agent in $\mathtt{A}$ agrees with this relative position between u and v (part 1), or else for every agent who does not agree there is a non-empty $\preccurlyeq$-cluster $\mathtt{B}$ containing agents with higher priority for whom v is preferred over u (part 2). ◀

The generalisation takes place in part 2. According to Definition 3.2, in order for $u \leq^{lr}_{\preccurlyeq} v$ to hold, for every agent i who does not consider v at least as plausible as u there should be an agent j with higher priority ($j \in [\prec\rangle(i)$) for whom v is preferred over u ($u <_j v$). But in Definition 3.3 the reliability ordering is just a preorder, so besides incomparable agents there might be also sets containing different but equally reliable agents. Thus, in order for $u \leq_{\preccurlyeq} v$ to hold, for every agent i who does not consider v at least as plausible as u there should be a *set of agents* $\mathtt{B}$, with all its members equally reliable ($j_1, j_2 \in \mathtt{B}$ implies $j_1 \simeq j_2$; part 2.1) and no other equally reliable agent being left out (for every $k \in \mathtt{A} \setminus \mathtt{B}$ there is a $j \in \mathtt{B}$ such that $j \not\simeq k$, from which it also follows that $\mathtt{B}$ is non-empty; part 2.2) —that is, a *cluster*— such that all agents in $\mathtt{B}$ have higher priority than i ($\mathtt{B} \subseteq [\prec\rangle(i)$) and *all of them agree* in that v is preferred over u ($u <_j v$ for all $j \in \mathtt{B}$; part 2.3). Of course, a different alternative for a generalisation is to stick with Definition 3.2, asking for the existence of just a single agent with higher priority who places v above u. This proposal's choice, to ask for the existence of a *cluster* with higher priority whose agents all place v above u, emphasises the fact that all agents in a cluster have the same priority, and thus in order for them to make a 'final' decision, they all should agree in their opinion.

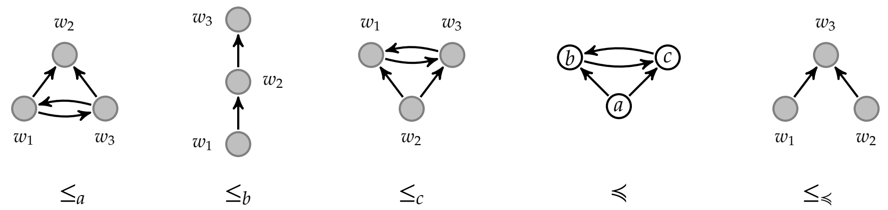
This lexicographic upgrade for preorders is indeed a generalisation of the lexicographic rule (for partial orders): when $\preccurlyeq$ is antisymmetric, the cluster $\mathtt{B}$ is forced to have at most one element (part 2.1) and to be non-empty (part 2.2), and therefore Definition 3.3 boils down to Definition 3.2. Formally,

Proposition 3.2 *Let $\{\leq_i\}_{i\in\mathtt{A}}$ be a finite collection of binary preference orderings over a domain W; let $\preccurlyeq$ be a partial order over $\mathtt{A}$. Then, for all $u, v \in W$,*

$$u \leq^{lr}_{\preccurlyeq} v \qquad \textit{iff} \qquad u \leq_{\preccurlyeq} v$$

Proof. From left to right, if i satisfies part 1 of Definition 3.2, then it satisfies part 1 of Definition 3.3, and if it satisfies part 2 of Definition 3.2, then the required j defines a set $\{j\}$ which, by $\preccurlyeq$'s reflexivity and antisymmetry, is a $\preccurlyeq$-cluster satisfying part 2 of Definition 3.3. Hence, $u \leq_{\preccurlyeq} v$. From right to left, if i satisfies part 1 of Definition 3.3, then it satisfies part 1 of Definition 3.2, and if it satisfies part 2 of Definition 3.3, then from $\preccurlyeq$'s antisymmetry the required $\preccurlyeq$-cluster $\mathtt{B}$ must be a singleton whose single element satisfies part 2 of Definition 3.2. Hence, $u \leq^{lr}_{\preccurlyeq} v$. For the formal details, see the Appendix (page 37). ■

Example 3.3 Consider the preference relations $\leq_a, \leq_b$ and $\leq_c$ from Example 3.1, shown below to the right. Consider the application of the lexicographic upgrade of Definition 3.3 with the reliability ordering shown below to the left.



The resulting preference ordering, $\leq_{\preccurlyeq}$, appears above on the right. Again, there is an edge from w_1 to w_3 because all preference orderings agree with this. On the other hand, there is an edge from w_2 to w_3 because b and c agree with this and, moreover, the fact that a does not agree is overruled by the opinion of all the members of the $\preccurlyeq$-cluster $\{b, c\}$, which lies $\prec$-above a. ◀

Some properties Here are some interesting properties of this lexicographic upgrade (and thus all the others it generalises). The first is that it respects unanimity about the relative position of any pair of worlds.

Proposition 3.3 *Let $\{\leq_i\}_{i\in\mathtt{A}}$ be a finite collection of binary preference orderings over a domain W, and let $\preccurlyeq$ be a preorder over $\mathtt{A}$. Take any u, v in W.*

(i) *If $u \simeq_i v$ for all $i \in \mathtt{A}$, then $u \simeq_{\preccurlyeq} v$.*

(ii) *If $u <_i v$ for all $i \in \mathtt{A}$, then $u <_{\preccurlyeq} v$.*

(iii) *If $u \nsim_i v$ for all $i \in \mathtt{A}$, then $u \nsim_{\preccurlyeq} v$.* ■

On the other hand, unanimity on the 'comparability' relation $\sim$ is not preserved: even if u and v are comparable for all agents in $\mathtt{A}$, the reliability ordering $\preccurlyeq$ might divide $\mathtt{A}$ into two disconnected regions with different opinions with respect to u and v's relative position (e.g., one of them a singleton whose element prefers u over v, and the other a singleton whose element prefers v over u). In such case, neither $u \leq_{\preccurlyeq} v$ nor $v \leq_{\preccurlyeq} u$ will hold. Still, comparability can be preserved when the reliability ordering satisfies extra requirements, as shown below when discussing the preservation of totality.

More important for this paper's purposes, the lexicographic upgrade preserves reflexivity and transitivity from the preference orderings.

Proposition 3.4 *Let $\{\leq_i\}_{i\in\mathtt{A}}$ be a finite collection of binary preference orderings over a domain W; let $\preccurlyeq$ be a preorder over $\mathtt{A}$. If every $\leq_i$ is reflexive (transitive), then so is $\leq_{\preccurlyeq}$.* ■

What about antisymmetry and totality, the two other properties typically required for relations representing orderings? For antisymmetry, the answer is straightforward: thanks to $\mathtt{A}$'s finiteness, such property is indeed preserved.

Proposition 3.5 *Let $\{\leq_i\}_{i\in\mathtt{A}}$ be a finite collection of binary preference orderings over a domain W; let $\preccurlyeq$ be a preorder over $\mathtt{A}$. If every $\leq_i$ is antisymmetric, then so is $\leq_{\preccurlyeq}$.* ■

The fact that antisymmetry is preserved relies on $\mathtt{A}$'s converse well-foundedness, a consequence of its finiteness: if an infinite $\prec$-ascending chain of $\preccurlyeq$-clusters were possible (e.g., $i_1 \prec i_2 \prec \ldots$), then while for every $i \in \mathtt{A}$ with $u <_i v$ there might be an i' strictly $\prec$-above i such that $v <_{i'} u$, for every $i' \in \mathtt{A}$ with $v <_{i'} u$ there might be an i strictly $\prec$-above i' such that $u <_i v$ (e.g., $u <_{i_\ell} v$ for ℓ odd, and $v <_{i_\ell} u$ for ℓ even). In such case, $u \leq_{\preccurlyeq} v$ and $v \leq_{\preccurlyeq} u$ would both hold, even if u and v are different.

For totality, the answer is more elaborate. First, as it has been mentioned, the lexicographic upgrade does not respect unanimity on comparability; thus, it does not preserve totality. Moreover, assuming that $\preccurlyeq$ is a *total* preorder is not enough. Take any pair u, v, and suppose not only that the single element of every singleton cluster considers u and v equally preferable, but also that all remaining (i.e. all non-singleton) clusters are 'not unanimous', i.e., each contains at least one element preferring u over v and one preferring v over u. If there is at least one non-singleton cluster, then neither $u \leq_{\preccurlyeq} v$ nor $v \leq_{\preccurlyeq} u$ holds: there is at least one agent $i \in \mathtt{A}$ (in the non-singleton cluster) for which $u \nleq_i v$ and there is no cluster whose elements all agree on $u < v$, so $u \nleq_{\preccurlyeq} v$, and there is at least one agent $i \in \mathtt{A}$ (in the non-singleton cluster) for which $v \nleq_i u$ and there is no cluster whose elements all agree on $v < u$, so $v \nleq_{\preccurlyeq} u$.

When looking for conditions guaranteeing that totality is preserved, the work of Andréka et al. [2] is helpful once again: as it is shown in its Theorem 4.1, the lexicographic rule of Definition 3.2 preserves totality when the reliability 'priority' relation is also total. Hence, as such upgrade is the particular case of the lexicographic upgrade in which $\preccurlyeq$ is antisymmetric (Proposition 3.2), it follows that the lexicographic upgrade of Definition 3.3 preserves totality when $\preccurlyeq$ is *both total and antisymmetric*. Indeed, when $\preccurlyeq$ is antisymmetric, every cluster is a singleton, and therefore it has a single 'unanimous' opinion about u and v's relative position. This forbids situations as those described in the previous paragraph, as if the single element of every singleton cluster (i.e., every agent) considers u and v equally preferable, there is unanimity on both $u \leq v$ and $v \leq u$, and thus both $u \leq_{\preccurlyeq} v$ and $v \leq_{\preccurlyeq} u$ hold. If, on the other hand, there is a cluster whose element prefers either u or else v above the other, simply pick, among such clusters, the one with the highest priority (which exists, as $\preccurlyeq$ is both total and antisymmetric): its single element will make the final decision if unanimity fails, making then either $u \leq_{\preccurlyeq} v$ or else $v \leq_{\preccurlyeq} u$ true.

Finally, below there is a slightly different (but clearly equivalent) alternative for formulating the definition of the lexicographic upgrade for preorders; it will be useful when providing an axiom system for the operation's associated dynamic modality. First, a couple of useful abbreviations.

Definition 3.4 Let $\{\leq_i\}_{i\in\mathtt{A}}$ be a finite collection of binary preference orderings over a domain W; let $\preccurlyeq$ be a preorder over $\mathtt{A}$ and take $\mathtt{B} \subseteq \mathtt{A}$.

- The set $C^i_{\preccurlyeq}$ contains those non-empty $\preccurlyeq$-clusters in $\mathtt{A}$ that lie $\prec$-above $i \in \mathtt{A}$:

$$C^i_{\preccurlyeq} := \{\mathtt{B} \subseteq \mathtt{A} \mid \mathtt{B} \text{ is a } \preccurlyeq\text{-cluster with } \mathtt{B} \neq \varnothing \text{ and } j \in \mathtt{B} \text{ implies } i \prec j\}$$

- The relation $<_{\mathtt{B}} \subseteq (W \times W)$ is the intersection of all the relations $<_j$ with $j \in \mathtt{B}$:

$$<_{\mathtt{B}} := \bigcap_{j\in\mathtt{B}} <_j$$

  Note how this is *not* equivalent to define $\leq_{\mathtt{B}}$ as $\bigcap_{j\in\mathtt{B}} \leq_j$ and then, following Definition 2.3, define $<_{\mathtt{B}}$ as $\leq_{\mathtt{B}} \cap \not\geq_{\mathtt{B}}$.[14] ◀

Then, the definition.

Definition 3.5 (Lexicographic upgrade, alternative definition) Let $\{\leq_i\}_{i\in\mathtt{A}}$ be a finite collection of binary preference orderings over a domain W; let $\preccurlyeq$ be a preorder over $\mathtt{A}$. The relation $\leq_{\preccurlyeq} \subseteq (W \times W)$ is defined as

$$\leq_{\preccurlyeq} := \bigcap_{i\in\mathtt{A}} \Big(\leq_i \cup \bigcup_{\mathtt{B}\in C^i_{\preccurlyeq}} <_{\mathtt{B}}\Big)$$

◀

Frame and model operations When the reliability ordering $\preccurlyeq$ is a preorder and the preference orderings in $\{\leq_i\}_{i\in\mathtt{A}}$ are preorders, the resulting (preference) ordering $\leq_{\preccurlyeq}$ is also a preorder (Proposition 3.4). Thus, consider any *PR* frame $\langle W, \{\leq_i, \preccurlyeq_i\}_{i\in\mathtt{A}}\rangle$: if each agent i upgrades her own preference ordering $\leq_i$ using her own reliability ordering $\preccurlyeq_i$, the resulting structure $\langle W, \{\leq'_i, \preccurlyeq_i\}_{i\in\mathtt{A}}\rangle$, with each $\leq'_i$ given by $\leq_{\preccurlyeq_i}$, is a *PR* frame too, as both its reliability and its preference relations are preorders. Here is the formal definition of the new structure.

Definition 3.6 (Frame and model operations) Let $F = \langle W, \{\leq_i, \preccurlyeq_i\}_{i\in\mathtt{A}}\rangle$ be a *PR* frame; let $M = \langle F, V\rangle$ be a *PR* model. The *PR* frame F_{lx}, differing from F only in the agents' preference relations, is the result of upgrading the preferences of *all* agents according to the lexicographic upgrade (Definition 3.3) and each agent's reliability relation. Formally, $F_{\mathrm{lx}} = \langle W, \{\leq'_i, \preccurlyeq_i\}_{i\in\mathtt{A}}\rangle$ is such that, for every $i \in \mathtt{A}$,

$$\leq'_i := \leq_{\preccurlyeq_i}$$

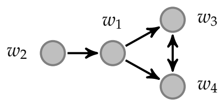
Accordingly, the model M_{lx} is formally defined as $M_{\mathrm{lx}} = \langle F_{\mathrm{lx}}, V\rangle$. ◀

Naturally, it is also possible to define frame/model operation in which the lexicographic upgrade is applied to the preference ordering of a single given agent (but see Subsection 4.1 for an alternative representation of such situations).

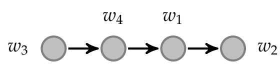
[14] For a simple counterexample, suppose the described definitions and take $W := \{w, u\}$ with $\leq_1 := W \times W$ and $\leq_2 := \leq_1 \setminus \{(u, w)\}$. Then, $<_1 = \varnothing$ and $<_2 = \{(w, u)\}$, and therefore $<_1 \cap <_2 = \varnothing$, but nevertheless $\leq_1 \cap \leq_2 = \leq_2$ and thus $(\leq_1 \cap \leq_2)$'s strict relation is $<_2$, that is, $\{(w, u)\}$.

Caveat The lexicographic policy is not the only 'reasonable' policy an agent can use for upgrading her preferences. For example, an agent a can upgrade her preferences by placing her most reliable agent's most preferred worlds above the rest, then using her old ordering within each zone. This upgrade, called *lexicographic* in Rott [98]; van Benthem [15] and *conservative* in Ghosh and Velázquez-Quesada [52], has been shown in Ghosh and Velázquez-Quesada [53] *not* to be an instance of their *glu* (which is 'equivalent' to the lexicographic upgrade of Definition 3.1 for totally ordered reliability relations), as there are cases in which the output of the former cannot be reproduced by any instance of the latter. The example below shows how this conservative upgrade is neither an instance of the more general lexicographic upgrade (Definition 3.3) that is central to this proposal.

Example 3.4 Suppose agent a is agent b's most reliable agent, $b \prec_b a$, and their individual preferences are as below (reflexive and transitive arrows omitted).

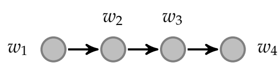


$\leq_a$



$\leq_b$

A conservative upgrade on b's preferences will create two zones, the upper one with a's most preferred worlds (w_3 and w_4), and the lower one with the remaining worlds (w_1 and w_2). Within each zone, b's old preferences will apply, thus placing w_4 strictly above w_4 and w_2 strictly above w_2. The resulting ordering $\leq'_b$ is, then,



$\leq'_b$

Clearly, a direct application of the lexicographic upgrade will produce a different outcome, as $w_2 \leq_{\preccurlyeq_b} w_1$ because not only $w_2 \leq_a w_1$ but also the $\preccurlyeq_b$-cluster $\{a\}$ lies $\prec_b$-above b and satisfies $w_2 <_a w_1$. Moreover, $\leq'$ cannot be recreated, even if the reliability ordering changes: ***(i)*** both $a \not\approxeq_b b$ and $a \approxeq_b b$ fail as, in such case, $w_1 \not\leq_{\preccurlyeq_b} w_2$ because agent a is such that $w_1 \not\leq_a w_2$ and there are no cluster of agents $\prec_b$-above a; ***(ii)*** $a \prec_b b$ fails as, in such case, $w_4 \leq_{\preccurlyeq_b} w_1$ because not only there is a $\preccurlyeq_b$-cluster $\{b\}$ lying $\prec_b$-above a and satisfying $w_4 <_b w_1$, but also $w_4 \leq_b w_1$. ◀

The conservative upgrade is not lexicographic: it does not create a preference ordering following a priority list of orderings. Instead, it uses a partition of the domain to create an ordered collection of 'clusters', using then a default ordering to sort the worlds within each 'cluster'. When working with preference relations that are total preorders and reliability relations that are total orders, this conservative upgrade and the *glu* are shown to be instances of the *general layered upgrade* [52]. Nevertheless, such operation lacks some desirable properties, the most important one being the preservation of unanimous preferences. Thus, under such upgrade policy, a reached 'full agreement' can be broken when 'further discussion' is allowed.

3.4 Coda: reliability as a total preorder

For the sake of completeness, this subsection provides an alternative definition of the lexicographic upgrade for those cases in which the reliability ordering is a *total* preorder (i.e., a total, reflexive and transitive relation), and thus although there might be different agents with the same priority, each one of them is comparable to any other.

Proposition 3.6 *Let $\{\leq_i\}_{i\in\mathtt{A}}$ be a finite collection of binary preference orderings over a domain W; let $\preccurlyeq$ be a* total preorder *over $\mathtt{A}$. Define the sets*

$$C_{\preccurlyeq} := \{\mathtt{B} \subseteq \mathtt{A} \mid \mathtt{B} \textit{ is a } \preccurlyeq\textit{-cluster with } \mathtt{B} \neq \varnothing\}$$
$$C_{\preccurlyeq}^{\mathtt{B}} := \{\mathtt{B}' \in C_{\preccurlyeq} \mid \textit{ for all } j \in \mathtt{B} \textit{ and } j' \in \mathtt{B}',\ j \prec j'\}$$

so $C_{\preccurlyeq}$ is the set of non-empty $\preccurlyeq$-clusters in $\mathtt{A}$ and, given $\mathtt{B} \subseteq \mathtt{A}$, $C_{\preccurlyeq}^{\mathtt{B}}$ is the set of all $\preccurlyeq$-clusters in $\mathtt{A}$ lying $\prec$-above all elements of $\mathtt{B}$. Then,

$$u \leq_{\preccurlyeq} v \qquad \textit{iff} \qquad \underbrace{\bigwedge_{i\in\mathtt{A}} u \leq_i v}_{1} \ \vee\ \underbrace{\bigvee_{\mathtt{B}\in C_{\preccurlyeq}} \Big(\bigwedge_{j\in\mathtt{B}} u <_j v \ \wedge\ \bigwedge_{\mathtt{B}'\in C_{\preccurlyeq}^{\mathtt{B}}} \bigwedge_{j'\in\mathtt{B}'} u \leq_{j'} v \Big)}_{2}$$

Proof. The statement says that $u \leq_{\preccurlyeq} v$ if and only if either all agents in $\mathtt{A}$ consider v at least as preferable as u, or else there is a non-empty cluster of agents $\mathtt{B}$ whose elements place v above u, with every element of every cluster with higher priority considering v at least as preferable as u. From left to right, if all agents satisfy part 1 of $\leq_{\preccurlyeq}$'s definition, then part 1 of the statement's right-hand side holds. Otherwise, for each i with $u \not\leq_i v$ there is a non-empty $\preccurlyeq$-cluster lying $\prec$-above i whose agents put all v above u. Since $\preccurlyeq$ is total and $\mathtt{A}$ is finite, among all such clusters it is possible to select the one with higher priority, $\mathtt{B}$, which satisfies the requirements on part 2 of the statement's right-hand side. From right to left, if part 1 holds, then clearly, $u \leq_{\preccurlyeq} v$. Otherwise, part 2 holds; let $\mathtt{B}$ be such $\preccurlyeq$-cluster, and consider any agent $i \in \mathtt{A}$. If i is in $\mathtt{B}$ or in a cluster above $\mathtt{B}$, then it is satisfies part 1 of $\leq_{\preccurlyeq}$'s definition; otherwise, it is in a cluster below $\mathtt{B}$, but then $\mathtt{B}$ itself is the needed $\preccurlyeq$-cluster required by part 2 of $\leq_{\preccurlyeq}$'s definition. Thus, $u \leq_{\preccurlyeq} v$. For the formal details, see the Appendix (page 40). ■

The advantage of this definition is that it is not given as a disjunction that each agent in $\mathtt{A}$ should satisfy, but rather as a disjunction that asks for either unanimity or otherwise the existence of a cluster that, when there is no unanimity, has the 'last word'.

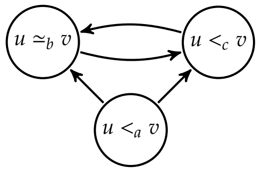
Addendum If the reader started by looking at the definition of the lexicographic upgrade for the total order case (Definition 3.1), and then asked to provide a definition for the total preorder case, she/he might have suggested the following:

$$\underbrace{\Big(\bigwedge_{i\in\mathrm{Mnm}_{\preccurlyeq}} u \leq_i v \ \wedge \bigwedge_{i\in\mathtt{A}\setminus\mathrm{Mnm}_{\preccurlyeq}} u \simeq_i v \Big)}_{1} \ \vee\ \underbrace{\bigvee_{\mathtt{B}\in C_{\preccurlyeq}^{\mathrm{Mnm}_{\preccurlyeq}}} \Big(\bigwedge_{j\in\mathtt{B}} u <_j v \ \wedge\ \bigwedge_{\mathtt{B}'\in C_{\preccurlyeq}^{\mathtt{B}}} \bigwedge_{j'\in\mathtt{B}'} u \simeq_{j'} v \Big)}_{2}$$

This alternative asks for either all non-minimum agents to consider u and v equally preferred and all minimum agents to consider v at least as preferred as u

(part 1), or else the existence of a cluster above the bottom whose elements have strict preference for v over u, and with all agents above such cluster seeing both worlds as equally preferred. This alternative indeed is closer to the definition for the total order case, but nevertheless it is not equivalent to the definition for the total preorder case provided above (Proposition 3.6).

Fact 3.1 *Consider three agents, a (total preorder) reliability ordering and two worlds* u, v *such that*



Under the equivalence stated in Proposition 3.6, $u \leq_{\preccurlyeq} v$ *is the case, as* $u \leq v$ *holds for all agents, and therefore part 1 holds. However, under the alternative provided just above,* $u \not\leq_{\preccurlyeq} v$*: part 1 fails because* $u \simeq v$ *is not the case for all non-minimum agents (in particular, it is not the case for* c*), and part 2 fails because the unique cluster whose elements all agree on* $u < v$*,* $\{a\}$*, is not above the bottom.* ■

In order to confirm that Proposition 3.6 truly provides the total preorder version of the lexicographic upgrades that have been defined so far, the following proposition shows how, when the reliability ordering is a total order, the 'definition' on Proposition 3.6 (*TP*) and Definition 3.1 (*TO*) are, indeed, equivalent.

Proposition 3.7 *Let* $\{\leq_i\}_{i \in \mathtt{A}}$ *be a finite collection of binary preference orderings over a domain* W*. Let* $\preccurlyeq$ *be a* total order *over* $\mathtt{A}$ *and thus, by Propositions 3.2 (equivalence for the partial order case) and 3.1 (equivalence for the total order case), take* $\leq_{\preccurlyeq}$ *as in Definition 3.1. Then,*

$$u \leq_{\preccurlyeq} v \qquad \textit{iff} \qquad \underbrace{\bigwedge_{i \in \mathtt{A}} u \leq_i v}_{TP-1} \;\vee\; \underbrace{\bigvee_{\mathtt{B} \in \mathcal{C}_{\preccurlyeq}} \Big(\bigwedge_{j \in \mathtt{B}} u <_j v \;\wedge\; \bigwedge_{\mathtt{B}' \in \mathcal{C}_{\preccurlyeq}^{\mathtt{B}}} \bigwedge_{j' \in \mathtt{B}'} u \leq_{j'} v \Big)}_{TP-2}$$

Proof. From right to left. Suppose *TP-1* holds, so all agents are such that $u \leq v$. If, additionally, all agents are such that $v \leq u$, then all agents are such that $u \simeq v$, and hence part 1 of Definition 3.1 (*TO-1*) holds. Otherwise, let j be the $\preccurlyeq$-maximum (recall: $\preccurlyeq$ is a total order) among those elements in $\mathtt{A}$ with $v \not\leq u$, so all elements above j are such that $v \leq u$, and hence $u \simeq v$. If j is also the $\preccurlyeq$-minimum, then *TO-1* holds; if it is above the minimum, part 2 of Definition 3.1 (*TO-2*) holds. Now suppose *TP-2* holds, so there is at least one cluster $\mathtt{B}$ whose elements all have $u < v$, and every element above $\mathtt{B}$ has $u \leq v$. Let $\mathtt{B}'$ be, among such clusters, the $\preccurlyeq$-maximum one, and as $\preccurlyeq$ is a total order, let $\mathtt{B}' = \{j'\}$. Since j' is the $\preccurlyeq$-maximum with $u < v$, all elements above are such that $u \not< v$, but as all in $\mathtt{A}$ satisfy $u \leq v$, all elements above j are such that $u \simeq v$. Again, if j' is the minimum, *TO-1* holds; otherwise, *TO-2* holds.

From left to right. Suppose *TO-1* holds; then clearly *TP-1* holds. Now suppose *TO-2* holds; then, as $\preccurlyeq$ is a total order, the witness j is such that the set $\{j\}$ is a cluster whose members satisfy all $u < v$, and all elements above clearly satisfy $u \leq v$. Hence, *TP-2* holds. ■

3.5 The formal language

In order to describe the changes the lexicographic upgrade operation brings about, the language is extended in the following way.

Definition 3.7 The language $\mathcal{L}_{\{\mathrm{lx}\}}$ extends $\mathcal{L}$ with the modality $\langle \mathrm{lx} \rangle$. Accordingly, given a *PR* model M, the function $[\![\cdot]\!]^M$ of Definition 2.6 is extended with the clause

$$[\![\langle \mathrm{lx} \rangle\, \varphi]\!]^M := [\![\varphi]\!]^{M_{\mathrm{lx}}}$$

thus stating that the worlds in the *PR* model M where $\langle \mathrm{lx} \rangle\, \varphi$ holds are exactly those worlds in the *PR* model M_{lx} (Definition 3.6) where φ holds.

Since the lexicographic upgrade on models is a total function and the semantic interpretation of its correspondent modality lacks a precondition (the agents can upgrade their preferences without any requirement), the set of worlds in M where the dual modality $[\mathrm{lx}]\, \varphi := \neg \langle \mathrm{lx} \rangle \neg\varphi$ holds is also given by $[\![\varphi]\!]^{M_{\mathrm{lx}}}$. Hence, $\langle \mathrm{lx} \rangle\, \varphi \leftrightarrow [\mathrm{lx}]\, \varphi$ is valid. ◀

With the modality $\langle \mathrm{lx} \rangle$ it is possible to express the effect of the simultaneous application of the lexicographic upgrade to every agent's preferences, with each agent using her own reliability ordering. Indeed, formulas of the form $\langle \mathrm{lx} \rangle\, \varphi$ express that after *every* agent performs a lexicographic upgrade, φ will be the case. If the reader is interested in a modality describing the effect of a lexicographic upgrade over the preferences of a *single* agent, it is enough to define such modality using the single-agent upgrade model operation sketched below Definition 3.6. (See also the variations of Section 4.)

Axiom system When looking for an axiom system for a modality describing the effect of a model operation (a *dynamic* modality), a useful *dynamic epistemic logic* [36; 16] strategy is to provide *recursion axioms*: valid formulas and validity-preserving rules indicating how to translate a formula with occurrences of the model-changing modality (a formula in the 'dynamic' language) into a provably equivalent one without them (a formula in the 'basic' 'static' language). Then, while soundness follows from the validity and validity-preserving properties of the new axioms and rules (so a formula and its translation are semantically equivalent), completeness follows from the completeness of the axiom system for the basic language, as the recursion axioms define a recursive validity-preserving translation from the 'dynamic' language into the 'basic' one. The reader is referred to Chapter 7 of van Ditmarsch et al. [36] (cf. Wang and Cao [103]) for an extensive explanation of this technique. Note that the existence of such recursion axioms (and hence of such translation) indicates that the 'basic' language is expressive enough to describe the changes the model operation induces: any formula in the 'dynamic' language (potentially expressing what happens after the model operation) is equivalent to a formula in the 'static' language (allowed to talk only about the current model).

In this proposal's case, the goal is then to provide recursion axioms defining a (validity-preserving) translation that takes a formula of the form $\langle \mathrm{lx} \rangle\, \varphi$ with φ in $\mathcal{L}$, and returns a formula $t(\langle \mathrm{lx} \rangle\, \varphi)$ in $\mathcal{L}$.[15] This is typically done in an inductive way, providing a recursion axiom for each possible φ.

[15] Cases with nested occurrences of the dynamic modality, as $\langle \mathrm{lx} \rangle \langle \mathrm{lx} \rangle\, p$, can be dealt with in a deepest-first fashion, as one can work with the subformula containing the deepest occurrence

The language $\mathcal{L}$ is based on the always true formula $\top$, atomic propositions p and formulas describing the agents' reliability relations $j \sqsubseteq_i k$, and it is closed under negations $\neg$, disjunctions $\vee$ and modal operators of the form $\langle \pi \rangle$ with π a relational expression. Hence, each one of these components requires an appropriate recursion axiom. For an operation as the lexicographic upgrade, a total function without precondition that affects neither atomic propositions nor reliability, the recursion axioms for $\top$, p, $j \sqsubseteq_i k$, $\neg$ and $\vee$ are standard (see Table 2). The interesting case is the one for modal operators $\langle \pi \rangle$: since π is an arbitrary relational expression, an appropriate translation must be presented in each case. In order to do this, it will be useful to have another translation: one taking a relational expression representing a relation in the upgraded model M_{lx}, and returning a relational expression representing a 'matching' relation in the original model M. The *relational transformer* defined below, similar in spirit to the program transformer of van Benthem et al. [17] for providing recursion axioms for regular *PDL*-expressions [69] (in their case, after action-model updates; Baltag et al. [8]), captures this. In order to simplify the presentation, here are some useful abbreviations.

Definition 3.8 Take any $a \in \mathtt{A}$ and any $\mathtt{B} \subseteq \mathtt{A}$.

- The formula $\text{nec}_a^{\mathtt{B}} \in \mathcal{L}^f$, expressing the fact that $\mathtt{B}$ is a non-empty $\preccurlyeq_a$-cluster, is defined as

$$\text{nec}_a^{\mathtt{B}} := \bigwedge_{j_1, j_2 \in \mathtt{B}} j_1 \equiv_a j_2 \;\;\wedge \bigwedge_{k \in (\mathtt{A} \setminus \mathtt{B})} \bigvee_{j \in \mathtt{B}} k \not\equiv_a j$$

- The formula $\text{abv}_a^{\mathtt{B},i} \in \mathcal{L}^f$, expressing the fact that the elements of $\mathtt{B}$ are all $\prec_a$-above a given $i \in \mathtt{A}$, is defined as

$$\text{abv}_a^{\mathtt{B},i} := \bigwedge_{j \in \mathtt{B}} i \sqsubset_a j$$

- The formula $\text{neca}_a^{\mathtt{B},i} \in \mathcal{L}^f$ is the conjunction of the two previous formulas:

$$\text{neca}_a^{\mathtt{B},i} := \text{nec}_a^{\mathtt{B}} \wedge \text{abv}_a^{\mathtt{B},i}$$

- The sets C_a^i, C_a and $C_a^{\mathtt{B}}$ are as in Definition 3.4 and Proposition 3.6, with a the subindex of the involved reliability relation (that is, $C_a^i := C_{\preccurlyeq_a}^i$, $C_a := C_{\preccurlyeq_a}$ and $C_a^{\mathtt{B}} := C_{\preccurlyeq_a}^{\mathtt{B}}$). ◀

Finally, here is the proper definition.

Definition 3.9 (Relational transformer) Define, for $a \in \mathtt{A}, \mathtt{B} \subseteq \mathtt{A}$ and $i \in \mathtt{A}$, the relational expressions

$$\tau_a^{\mathtt{B},i} := ?(\text{neca}_a^{\mathtt{B},i}, \text{neca}_a^{\mathtt{B},i}) \cap <_{\mathtt{B}} \qquad \text{and} \qquad \Upsilon_a^{\mathtt{B},i} := ?(\text{neca}_a^{\mathtt{B},i}, \text{neca}_a^{\mathtt{B},i}) \cap >_{\mathtt{B}}$$

of $\langle \text{lx} \rangle$ (in the example, $\langle \text{lx} \rangle p$), and afterwards (i.e., once such occurrence has been 'eliminated') proceed with the following one. Of course, for this to be allowed, the axiom system must include the rule of substitution of provably equivalents.

with

$$<_{\mathtt{B}} := \bigcap_{j \in \mathtt{B}} <_j \qquad \text{and} \qquad >_{\mathtt{B}} := \bigcap_{j \in \mathtt{B}} >_j$$

A *relational transformer* $Tx : \mathcal{L}^r \to \mathcal{L}^r$ is a function from relational expressions to relational expressions defined inductively as follows.

$$Tx(\leq_a) := \bigcap_{i \in \mathtt{A}} \Big(\leq_i \cup \bigcup_{\mathtt{B} \subseteq \mathtt{A}} \tau_a^{\mathtt{B},i} \Big) \qquad Tx(-\pi) := -Tx(\pi)$$

$$Tx(\geq_a) := \bigcap_{i \in \mathtt{A}} \Big(\geq_i \cup \bigcup_{\mathtt{B} \subseteq \mathtt{A}} \reflectbox{τ}_a^{\mathtt{B},i} \Big) \qquad Tx(\pi \cup \sigma) := Tx(\pi) \cup Tx(\sigma)$$

$$Tx(1) := 1 \qquad Tx(\pi \cap \sigma) := Tx(\pi) \cap Tx(\sigma)$$

$$Tx(?(\varphi, \psi)) := ?\,(\langle \mathrm{lx} \rangle\, \varphi, \langle \mathrm{lx} \rangle\, \psi) \qquad \blacktriangleleft$$

A relational transformer Tx takes a relational expression representing a relation in a model M_{lx} and returns a relational expression describing the same relation *in the original model* M. The cases for the basic relational expressions $\leq_a$ and $\geq_a$ are the important ones; thanks to the expressivity of the 'static' language $\mathcal{L}$, they use Definition 3.5 almost literally in order to indicate, first, that $\leq_a$ in M_{lx} corresponds to $\leq_{\preccurlyeq_a}$, the result of applying the lexicographic upgrade in M using $\preccurlyeq_a$ as the priority ordering among the preferences $\{\leq_i\}_{i \in \mathtt{A}}$, and second, that the relation $\geq_a$ in M_{lx} is simply the converse of $\leq_a$ in the same model, and hence the converse of $\leq_{\preccurlyeq_a}$ in M. The remaining cases take care of the constant 1, the relational test and the complement, union and intersection of relations.[16] The crucial property of the relational transformer is the following one.

Theorem 2 *Let $M = \langle W, \{\leq_i, \preccurlyeq_i\}_{i \in \mathtt{A}}, V \rangle$ be a PR model. Then, for every relational expression $\pi \in \mathcal{L}_r$,*

$$\langle\!\langle \pi \rangle\!\rangle^{M_{\mathrm{lx}}} = \langle\!\langle Tx(\pi) \rangle\!\rangle^{M}$$

Proof. The proof proceeds by induction on π's structure. Here, only the case for $\leq_a$ is discussed; for the rest, see the Appendix (page 40).

It will be shown that $\langle\!\langle \leq_a \rangle\!\rangle^{M_{\mathrm{lx}}} = \langle\!\langle Tx(\leq_a) \rangle\!\rangle^{M}$. By $\langle\!\langle \cdot \rangle\!\rangle$'s and M_{lx}'s definitions, $\langle\!\langle \leq_a \rangle\!\rangle^{M_{\mathrm{lx}}} = \leq_{\preccurlyeq_a}$, so the goal is to prove $\leq_{\preccurlyeq_a} = \langle\!\langle Tx(\leq_a) \rangle\!\rangle^{M}$. Now, for the left-hand side (Definition 3.5),

$$\leq_{\preccurlyeq_a} \;=\; \bigcap_{i \in \mathtt{A}} \Big(\leq_i \cup \bigcup_{\mathtt{B} \in C_a^i} <_{\mathtt{B}} \Big)$$

and, for the right-hand side,

$$\begin{aligned} \langle\!\langle Tx(\leq_a) \rangle\!\rangle^{M} &= \langle\!\langle \bigcap_{i \in \mathtt{A}} \Big(\leq_i \cup \bigcup_{\mathtt{B} \subseteq \mathtt{A}} \tau_a^{\mathtt{B},i} \Big) \rangle\!\rangle^{M} = \bigcap_{i \in \mathtt{A}} \langle\!\langle \leq_i \cup \bigcup_{\mathtt{B} \subseteq \mathtt{A}} \tau_a^{\mathtt{B},i} \rangle\!\rangle^{M} \\ &= \bigcap_{i \in \mathtt{A}} \Big(\leq_i \cup \langle\!\langle \bigcup_{\mathtt{B} \subseteq \mathtt{A}} \tau_a^{\mathtt{B},i} \rangle\!\rangle^{M} \Big) \end{aligned}$$

[16] The relational transformers presented here are more complex than those presented in Ghosh and Velázquez-Quesada [53], as they require not only reliability formulas to describe properties of sets of agents, but also the test operator ? to relate potentially disconnected worlds.

Then, it is enough to show that $\bigcup_{\mathtt{B}\in C_a^i} <_{\mathtt{B}}$ and $\langle\!\langle \bigcup_{\mathtt{B}\subseteq\mathtt{A}} \tau_a^{\mathtt{B},i} \rangle\!\rangle^M$ are the same. For this, consider the second, $\langle\!\langle \bigcup_{\mathtt{B}\subseteq\mathtt{A}} \tau_a^{\mathtt{B},i} \rangle\!\rangle^M$: it contains the union of all $\tau_a^{\mathtt{B},i}$ for $\mathtt{B} \subseteq \mathtt{A}$, with $\tau_a^{\mathtt{B},i} = ?(\mathrm{neca}_a^{\mathtt{B},i}, \mathrm{neca}_a^{\mathtt{B},i}) \cap \boldsymbol{<}_{\mathtt{B}}$. But recall that the formula $\mathrm{neca}_a^{\mathtt{B},i}$ expresses the fact that $\mathtt{B}$ is a non-empty $\preccurlyeq_a$-cluster that lies $<_a$-above i, that is, that $\mathtt{B} \in C_a^i$. Moreover, the formula only talks about the reliability relation, so it is either globally true or else globally false. Hence,

$$\mathtt{B} \in C_a^i \ \text{ implies } \ [\![\mathrm{neca}_a^{\mathtt{B},i}]\!]^M = W \qquad \text{and} \qquad \mathtt{B} \notin C_a^i \ \text{ implies } \ [\![\mathrm{neca}_a^{\mathtt{B},i}]\!]^M = \varnothing$$

and therefore

$$\begin{aligned} \mathtt{B} \in C_a^i \ &\text{ implies } \ \langle\!\langle ?(\mathrm{neca}_a^{\mathtt{B},i}, \mathrm{neca}_a^{\mathtt{B},i}) \rangle\!\rangle^M = W \times W \\ \mathtt{B} \notin C_a^i \ &\text{ implies } \ \langle\!\langle ?(\mathrm{neca}_a^{\mathtt{B},i}, \mathrm{neca}_a^{\mathtt{B},i}) \rangle\!\rangle^M = \varnothing \end{aligned}$$

Then, as $\langle\!\langle \tau_a^{\mathtt{B},i} \rangle\!\rangle^M = \langle\!\langle ?(\mathrm{neca}_a^{\mathtt{B},i}, \mathrm{neca}_a^{\mathtt{B},i}) \rangle\!\rangle^M \cap \langle\!\langle \boldsymbol{<}_{\mathtt{B}} \rangle\!\rangle^M$,

$$\langle\!\langle \tau_a^{\mathtt{B},i} \rangle\!\rangle^M = \begin{cases} \langle\!\langle \boldsymbol{<}_{\mathtt{B}} \rangle\!\rangle^M & \text{if } \mathtt{B} \in C_a^i \\ \varnothing & \text{if } \mathtt{B} \notin C_a^i \end{cases}$$

Thus, $\langle\!\langle \bigcup_{\mathtt{B}\subseteq\mathtt{A}} \tau_a^{\mathtt{B},i} \rangle\!\rangle^M = \bigcup_{\mathtt{B}\in C_a^i} \langle\!\langle \boldsymbol{<}_{\mathtt{B}} \rangle\!\rangle^M$. Finally, since $\boldsymbol{<}_{\mathtt{B}} = \bigcap_{j\in\mathtt{B}} \boldsymbol{<}_j$, it follows that $\langle\!\langle \boldsymbol{<}_{\mathtt{B}} \rangle\!\rangle^M = <_{\mathtt{B}}$; thus, $\langle\!\langle \bigcup_{\mathtt{B}\subseteq\mathtt{A}} \tau_a^{\mathtt{B},i} \rangle\!\rangle^M = \bigcup_{\mathtt{B}\in C_a^i} <_{\mathtt{B}}$, and the proof is complete. ■

With Tx defined, here are the recursion axioms.

$\vdash \langle \mathrm{lx} \rangle \top \ \leftrightarrow \ \top$	$\vdash \langle \mathrm{lx} \rangle \neg\varphi \ \leftrightarrow \ \neg \langle \mathrm{lx} \rangle \varphi$
$\vdash \langle \mathrm{lx} \rangle p \ \leftrightarrow \ p$	$\vdash \langle \mathrm{lx} \rangle (\varphi \vee \psi) \ \leftrightarrow \ (\langle \mathrm{lx} \rangle \varphi \vee \langle \mathrm{lx} \rangle \psi)$
$\vdash \langle \mathrm{lx} \rangle\, j \sqsubseteq_i j' \ \leftrightarrow \ j \sqsubseteq_i j'$	$\vdash \langle \mathrm{lx} \rangle \langle \pi \rangle \varphi \ \leftrightarrow \ \langle Tx(\pi) \rangle \langle \mathrm{lx} \rangle \varphi$
From $\vdash \varphi$ infer $\vdash [\mathrm{lx}]\, \varphi$	From $\vdash \psi_1 \leftrightarrow \psi_2$ infer $\vdash \varphi \leftrightarrow \varphi[\psi_2/\psi_1]$, with $\varphi[\psi_2/\psi_1]$ any formula obtained by replacing one or more occurrences of ψ_1 in φ with ψ_2.

Table 2: Recursion axioms for $\mathcal{L}_{\{\mathrm{lx}\}}$ w.r.t. *PR* models.

Theorem 3 *The axioms and rules on Table 1 together with those on Table 2 provide a sound and complete axiom system (with π any relational expression) for $\mathcal{L}_{\{\mathrm{lx}\}}$ with respect to PR models.*

Proof. (Sketch) Soundness follows from the validity and validity-preserving properties of the new axioms and rules. Here only the validity of the recursion axiom for relational expressions is discussed. Let $M = \langle W, \{\leq_i, \preccurlyeq_i\}_{i\in\mathtt{A}}, V \rangle$ be a *PR* model, and take any $w \in W$ such that $w \in [\![\langle \mathrm{lx} \rangle \langle \pi \rangle \varphi]\!]^M$. Then, from $[\![\cdot]\!]$'s definition, $w \in [\![\langle \pi \rangle \varphi]\!]^{M_{\mathrm{lx}}}$ and therefore $\langle\!\langle \pi \rangle\!\rangle_w^{M_{\mathrm{lx}}} \cap [\![\varphi]\!]^{M_{\mathrm{lx}}} \neq \varnothing$. But then, by Theorem 2 and Definition 3.7, $\langle\!\langle Tx(\pi) \rangle\!\rangle_w^M \cap [\![\langle \mathrm{lx} \rangle \varphi]\!]^M \neq \varnothing$, that is, $w \in [\![\langle Tx(\pi) \rangle \langle \mathrm{lx} \rangle \varphi]\!]^M$. The other direction is similar; then, $[\![\langle \mathrm{lx} \rangle \langle \pi \rangle \varphi]\!]^M = [\![\langle Tx(\pi) \rangle \langle \mathrm{lx} \rangle \varphi]\!]^M$, that is,

$$[\![\langle \mathrm{lx} \rangle \langle \pi \rangle \varphi \leftrightarrow \langle Tx(\pi) \rangle \langle \mathrm{lx} \rangle \varphi]\!]^M = W$$

Since M is an arbitrary model, $\langle \mathrm{lx} \rangle \langle \pi \rangle \varphi \leftrightarrow \langle Tx(\pi) \rangle \langle \mathrm{lx} \rangle \varphi$ is valid.

Completeness follows from the completeness of the basic system, as valid axioms and validity-preserving rules of Table 2 define a validity-preserving translation from formulas in $\mathcal{L}_{\{\mathrm{lx}\}}$ into formulas in $\mathcal{L}$.[17] ■

If the reader is interested in a modality describing the effect of a lexicographic upgrade over the preferences of a *single* given agent a, notice how such modality is also axiomatised by the presented system as long as the cases for $\leq_a$ and $\geq_a$ in the definition of the relational transformer Tx (Definition 3.9) are as they are *only for the given agent a*, with $Tx(\leq_i) := \leq_i$ and $Tx(\geq_i) := \geq_i$ for any other agent i.

4 An exploration of 'non-public' variations

The lexicographic upgrade represents the effect of a *public* announcement of the agents' individual preferences, leading each individual agent to upgrade her own preferences according to what has been announced and the reliability ordering she assigns to the set of agents. The 'public' aspect of this announcement is reflected in the fact that $\leq_{\preccurlyeq}$'s definition uses *all* agents in $\mathtt{A}$ (indicating that every agent can 'hear' every other agent), and the definition is the same for *every* pair of worlds u, v in the domain (indicating not only that every agent announces her full preference ordering, but also that every agent pays attention to every other agent's full preferences).

Of course, the communication among agents might not be always perfect. This section explores alternative definitions of the lexicographic upgrade for such cases.

4.1 Each agent can only 'hear' a subset of them

With respect to $\leq_{\preccurlyeq_a}$'s definition, this variation is straightforward, as it simply amounts to restrict the agents that are involved, from the full set $\mathtt{A}$, to the subset of agents the given a can actually 'hear', $\mathtt{A}_a$. Still, this is an interesting and useful case: as it has been mentioned, it allows the representation of natural upgrades that arise when the preferences of some agents in $\mathtt{A}$ are not used for building up the new preference ordering.

Definition 4.1 (Lexicographic upgrade, var 1) Let $M = \langle W, \{\leq_i, \preccurlyeq_i\}_{i \in \mathtt{A}}, V \rangle$ be a *PR* model; let $a \in \mathtt{A}$ be an agent, and let $\mathtt{A}_a \subseteq \mathtt{A}$ be a non-empty subset of agents. The relation $\leq'_{\preccurlyeq_a} \subseteq (W \times W)$ is defined as

$$\leq'_{\preccurlyeq_a} := \bigcap_{i \in \mathtt{A}_a} \Big(\leq_i \cup \bigcup_{\mathtt{B} \in C^i_a(\mathtt{A}_a)} <_{\mathtt{B}} \Big)$$

with $C^i_a(\mathtt{A}_a)$ given by

$$C^i_a(\mathtt{A}_a) := \{\mathtt{B} \subseteq \mathtt{A}_a \mid \mathtt{B} \text{ is a } \preccurlyeq_a\text{-cluster w.r.t. } \mathtt{A}_a \text{ s.t. } \mathtt{B} \neq \varnothing \text{ and } j \in \mathtt{B} \text{ implies } i \prec_a j\}$$

The only difference between this definition and Definition 3.5, besides the explicit use of agent a's reliability ordering $\preccurlyeq_a$, is that here $\mathtt{A}$ has been replaced

[17] The rule of substitution of logical equivalents, at the bottom right of the table, takes care of formulas with more than one occurrence of $\langle \mathrm{lx} \rangle$, as one can work with the deepest occurrence of such modality and, once it is eliminated, proceed with the following one.

by $\mathtt{A}_a$: only the preferences of agents in such set are involved in the definition of a's new preference ordering. ◀

After providing the upgrade's definition, two questions arise. First, does this upgrade preserve *PR* models? Second, is there a relational transformer for it, and thus a valid recursion axiom for its associated modality? The answer for the first is yes: the new definition is a simple variation of Definition 3.5 that uses $\mathtt{A}_a$ as the domain, and thus it preserves reflexivity and transitivity.

The answer for the second is also yes for those cases in which there is a formula characterising the set $\mathtt{A}_a$, that is, a formula ist_a^i in $\mathcal{L}^f$ such that, for every *PR* model M and agent $i \in \mathtt{A}$, $i \in \mathtt{A}_a$ implies $[\![\mathrm{ist}_a^i]\!]^M = W$ and $i \notin \mathtt{A}_a$ implies $[\![\mathrm{ist}_a^i]\!]^M = \varnothing$. In such cases, after extending ist_a^i to $\mathrm{ist}_a^{\mathtt{B}}$ in the natural way ($\mathrm{ist}_a^{\mathtt{B}} := \bigwedge_{j\in\mathtt{B}} \mathrm{ist}_a^j$), it is possible to define the formulas

$$\begin{aligned}
\widehat{\mathrm{nec}}_a^{\mathtt{B}} &:= \mathrm{ist}_a^{\mathtt{B}} \wedge \bigwedge_{j_1,j_2\in\mathtt{B}} j_1 \equiv_a j_2 \quad\wedge \bigwedge_{k\in(\mathtt{A}\setminus\mathtt{B})} \Big(\mathrm{ist}_a^k \rightarrow \bigvee_{j\in\mathtt{B}} k \not\equiv_a j\Big)\\
\widehat{\mathrm{neca}}_a^{\mathtt{B},i} &:= \widehat{\mathrm{nec}}_a^{\mathtt{B}} \wedge \mathrm{abv}_a^{\mathtt{B},i}
\end{aligned}$$

with the first expressing that $\mathtt{B}$ is a non-empty $\preccurlyeq_a$-cluster with respect to $\mathtt{A}_a$ (cf. with the definition of $\mathrm{nec}_a^{\mathtt{B}}$ in Definition 3.8) and the second expressing that, additionally, all elements of $\mathtt{B}$ are $\prec_a$-above i. Then, it is possible to define the relational expression

$$\widehat{\tau}_a^{\,\mathtt{B},i} := ?(\widehat{\mathrm{neca}}_a^{\mathtt{B},i}, \widehat{\mathrm{neca}}_a^{\mathtt{B},i}) \cap <_{\mathtt{B}}$$

Now, note how

- if $\mathtt{B} \in C_a^i(\mathtt{A}_a)$ then $[\![\widehat{\mathrm{neca}}_a^{\mathtt{B},i}]\!]^M = W$, thus $\langle\!\langle ?(\widehat{\mathrm{neca}}_a^{\mathtt{B},i}, \widehat{\mathrm{neca}}_a^{\mathtt{B},i})\rangle\!\rangle^M = W \times W$ and hence $\langle\!\langle \widehat{\tau}_a^{\,\mathtt{B},i}\rangle\!\rangle^M = <_{\mathtt{B}}$, and

- if $\mathtt{B} \notin C_a^i(\mathtt{A}_a)$ then $[\![\widehat{\mathrm{neca}}_a^{\mathtt{B},i}]\!]^M = \varnothing$, thus $\langle\!\langle ?(\widehat{\mathrm{neca}}_a^{\mathtt{B},i}, \widehat{\mathrm{neca}}_a^{\mathtt{B},i})\rangle\!\rangle^M = \varnothing$ and hence $\langle\!\langle \widehat{\tau}_a^{\,\mathtt{B},i}\rangle\!\rangle^M = \varnothing$.

Therefore,

$$\langle\!\langle \widehat{\tau}_a^{\,\mathtt{B},i}\rangle\!\rangle^M = \begin{cases} \langle\!\langle <_{\mathtt{B}}\rangle\!\rangle^M & \text{if } \mathtt{B} \in C_a^i(\mathtt{A}_a)\\ \varnothing & \text{if } \mathtt{B} \notin C_a^i(\mathtt{A}_a)\end{cases}$$

and thus $\langle\!\langle \bigcup_{\mathtt{B}\subseteq\mathtt{A}} \widehat{\tau}_a^{\,\mathtt{B},i}\rangle\!\rangle^M = \bigcup_{\mathtt{B}\in C_a^i(\mathtt{A}_a)} \langle\!\langle <_{\mathtt{B}}\rangle\!\rangle^M$. Moreover, it is also possible to define the relational expression

$$\widehat{\varsigma}_a^{\,\mathtt{B},i} := \Big(\leq_i \cup \bigcup_{\mathtt{B}\subseteq\mathtt{A}} \widehat{\tau}_a^{\,\mathtt{B},i}\Big) \cup ?(\neg\mathrm{ist}_a^i, \neg\mathrm{ist}_a^i)$$

Again, note how

- if $i \in \mathtt{A}_a$, then $[\![\mathrm{ist}_a^i]\!]^M = W$ so $[\![\neg\mathrm{ist}_a^i]\!]^M = \varnothing$, hence $\langle\!\langle ?(\neg\mathrm{ist}_a^i, \neg\mathrm{ist}_a^i)\rangle\!\rangle^M = \varnothing$ and therefore $\langle\!\langle \widehat{\varsigma}_a^{\,\mathtt{B},i}\rangle\!\rangle^M = \langle\!\langle \leq_i \cup \bigcup_{\mathtt{B}\subseteq\mathtt{A}} \widehat{\tau}_a^{\,\mathtt{B},i}\rangle\!\rangle^M$;

- if $i \notin \mathtt{A}_a$ then $[\![\mathrm{ist}_a^i]\!]^M = \varnothing$ so $[\![\neg\mathrm{ist}_a^i]\!]^M = W$, hence $\langle\!\langle ?(\neg\mathrm{ist}_a^i, \neg\mathrm{ist}_a^i)\rangle\!\rangle^M = W \times W$ and therefore $\langle\!\langle \widehat{\varsigma}_a^{\,\mathtt{B},i}\rangle\!\rangle^M = W \times W$.

Therefore,

$$\langle\!\langle \widehat{\varsigma}_a^{\mathtt{B},i} \rangle\!\rangle^M = \begin{cases} \langle\!\langle \leq_i \cup \bigcup_{\mathtt{B}\subseteq\mathtt{A}} \widehat{\tau}_a^{\mathtt{B},i} \rangle\!\rangle^M & \text{if } i \in \mathtt{A}_a \\ W \times W & \text{if } i \notin \mathtt{A}_a \end{cases}$$

and thus $\langle\!\langle \bigcap_{i\in\mathtt{A}} \widehat{\varsigma}_a^{\mathtt{B},i} \rangle\!\rangle^M = \bigcap_{i\in\mathtt{A}_a} \langle\!\langle \leq_i \cup \bigcup_{\mathtt{B}\subseteq\mathtt{A}} \widehat{\tau}_a^{\mathtt{B},i} \rangle\!\rangle^M$. With these abbreviations, the relational transformer for $\leq_a$ can be redefined simply as

$$Tx'(\leq_a) := \bigcap_{i\in\mathtt{A}} \widehat{\varsigma}_a^{\mathtt{B},i} = \bigcap_{i\in\mathtt{A}} \Big(\big(\leq_i \cup \bigcup_{\mathtt{B}\subseteq\mathtt{A}} \widehat{\tau}_a^{\mathtt{B},i} \big) \cup ?(\neg\mathrm{ist}_a^i, \neg\mathrm{ist}_a^i) \Big)$$

As a simple example, if agent a only listen to her most reliable agents, then it is enough to define $\mathrm{ist}_a^i := \bigwedge_{k\in\mathtt{A}} \neg(i \sqsubset_a k)$. Interestingly, by defining $\mathtt{A}_i = \{i, a\}$ for all $i \in \mathtt{A}$ and a fixed a, the operation of Definition 4.1 allows the representation of situations in which only one agent (the fixed a) announces her preferences, and all of them update their own based on the relative reliability they assign to the one announcing her preferences (a) and themselves. This allows the representation of some form of *sequential* announcement of the individual preferences.

4.2 Each agent only 'pays attention' to part of the announced preferences

Suppose agent a only cares about a subset U_a of the given domain W, indicating thus that she will keep her own ordering among those elements in $W \setminus U_a$.

Definition 4.2 (Lexicographic upgrade, var 2) Let $M = \langle W, \{\leq_i, \preccurlyeq_i\}_{i\in\mathtt{A}}, V\rangle$ be a *PR* model; let $a \in \mathtt{A}$ be an agent, and let $U_a \subseteq W$ be a subset of the domain. The relation $\leq'_{\preccurlyeq_a} \subseteq (W \times W)$ is defined as

$$\leq'_{\preccurlyeq_a} := \Big(\bigcap_{i\in\mathtt{A}} (\leq_i \cup \bigcup_{\mathtt{B}\in C_a^i} <_{\mathtt{B}}) \cap (U_a \times U_a) \Big) \cup \Big(\leq_a \cap\, \overline{U_a \times U_a} \Big)$$

with $\overline{U_a \times U_a} := (W \times W) \setminus (U_a \times U_a)$, as usual. Thus, $\leq'_{\preccurlyeq_a}$ is as in Definition 3.5 for those pairs in $U_a \times U_a$, and simply as $\leq_a$ for the rest. ◀

Although this operation preserves reflexivity, it does not preserve transitivity. An edge between two given worlds might exists for two different reasons (both worlds are in U_a and they satisfy Definition 3.5, or one of them is not in U_a and they are related by $\leq_a$), and if the reason for an edge from a given u to a given v is different from the one for an edge from v to a given w, then there might be no reason for an edge from u to w.[18] This does not imply that the given definition is of not use: it only makes the properties of the set U_a important.[19] A study of the conditions for U_a under which the operation preserves transitivity is left for future work.

---

[18] Consider agents a, b, with their (reflexive and transitive) preferences over $W = \{u, v, w\}$ given by $\leq_a := \mathrm{Id}_W \cup \{(v, w)\}$ and $\leq_b := \mathrm{Id}_W \cup \{(u, v)\}$. If $U_a := \{u, v\}$ and $a \prec_a b$, then $u \leq_{\preccurlyeq_a} v$ (as $u, v \in U_a$ and, while $u \leq_b v$, a $\prec_a$-sees a more reliable cluster $\{b\}$ with $u <_b v$) and $v \leq_{\preccurlyeq_a} w$ (as $w \notin U_a$ but $v \leq_a w$) but nevertheless $u \nleq_{\preccurlyeq_a} w$ (as $w \notin U_a$ and $u \nleq_a w$).

[19] For example, if the set U_a is the full domain W or the empty set $\varnothing$, then reflexivity and transitivity are clearly preserved.

With respect to an axiom system, the definition of a proper relational transformer is straightforward. If the formula χ_a characterises the set U_a (that is, for every PR model M, $U_a = [\![\chi_a]\!]^M$), then

$$Tx'(\leq_a) := \big(Tx(\leq_a) \cap ?(\chi_a, \chi_a)\big) \cup \big(\leq_a \cap ?(\top, \neg\chi_a)\big) \cup \big(\leq_a \cap ?(\neg\chi_a, \top)\big)$$

with Tx the original relational transformer of Definition 3.9.

It is worthwhile to notice how this variation, which allows agents to focus on subsets of the domain, also allows the representation of a different scenario: one in which agents either perform the lexicographic upgrade over the full domain, or else do nothing. Indeed, situations in which only agents in $\mathtt{B} \subseteq \mathtt{A}$ upgrade their preferences fully with the rest of them keeping theirs as before can be represented simply by taking W as the 'relevant set' for all agents $i \in \mathtt{B}$ (so they will upgrade their full preferences by means of the lexicographic upgrade) and $\emptyset$ as the 'relevant set' for agents $i \notin \mathtt{B}$ (so they will keep their old preference ordering). Interestingly, both reflexivity and transitivity are preserved with such definitions for the relevant sets.

4.3 Each agent can only 'announce' part of her preferences

Suppose each agent i can only announce her preferences about a subset U_i of the domain. In such case, the lexicographic uprgade can be defined as follows.

Definition 4.3 (Lexicographic upgrade, var 3) Let $M = \langle W, \{\leq_i, \preccurlyeq_i\}_{i\in\mathtt{A}}, V\rangle$ be a PR model; let $a \in \mathtt{A}$ be an agent, and let $\{U_i\}_{i\in\mathtt{A}}$ be a collection of set of worlds, one for every agent $i \in \mathtt{A}$. The relation $\leq'_{\preccurlyeq_a} \subseteq (W \times W)$ is defined as

$$\leq'_{\preccurlyeq_a} := \bigcap_{i\in\mathtt{A}} \Big(\big(\leq_i \cap (U_i \times U_i)\big) \cup \bigcup_{\mathtt{B}\in C_a^i} \bigcap_{j\in\mathtt{B}} \big(<_j \cap (U_j \times U_j)\big)\Big)$$

Thus, the use of each preference ordering $\leq_i$ is restricted to those pairs in $U_i \times U_i$, indicating that only what agent i can announce is relevant. (If agent a will use her own ordering $\leq_a$ for those pairs not in $U_i \times U_i$, it is enough to add an expression of the form $\leq_a \cap \overline{U_i \times U_i}$ in the appropriate place, analogous to the definition of the previous variation, Definition 4.2.) ◀

Just as in the case of Definition 4.2, this variation does not preserve transitivity;[20] again, additional conditions over $\{U_i\}_{i\in\mathtt{A}}$ that make the operation preserve PR models is left for future work. An appropriate relational transformer can be defined along the lines of the previous ones:

$$Tx'(\leq_a) := \bigcap_{i\in\mathtt{A}} \Big(\big(\leq_i \cap ?(\chi_i, \chi_i)\big) \cup \bigcup_{\mathtt{B}\in\mathtt{A}} \widehat{\tau}_a^{\,\mathtt{B},i}\Big)$$

with

$$\widehat{\tau}_a^{\,\mathtt{B},i} := ?(\mathrm{neca}_a^{\mathtt{B},i}, \mathrm{neca}_a^{\mathtt{B},i}) \cap \bigcap_{j\in\mathtt{B}} \big(<_j \cap ?(\chi_j, \chi_j)\big)$$

and each χ_i characterising its corresponding set U_i.

[20]It does not even preserve reflexivity, as the sets U_i are not required to be collectively exhaustive.

5 Related work

As mentioned in the introduction, the present proposal recreates collective decision making scenarios but, instead of aligning itself with either the *aggregation* or else the *deliberative* alternatives discussed already in the introduction, it combines ideas of both with some 'real life' limitations, producing a framework whose behaviour is closer to social psychology situations (opinion and information flow) in which the preferences/beliefs/behaviour of a group of agents influence the individual's preferences/beliefs/behaviour of an individual.

The usual approach to study this social dynamics phenomena is to use statistical modelling, and such studies can be traced back to French [48]. A classic proposal in this tradition is that of DeGroot [31] on reaching a consensus, which considers a set of individuals, each one with both a probability distribution (for the value of a given parameter) and a weight for the probability distribution of each individual. It proposes a model through which agents change their probability distribution after being informed of the probability distribution of one another's, the crucial part being the definition of each individual's revised distribution as a linear combination of the distributions of the group's members. The present proposal can indeed be understood as a qualitative version of DeGroot [31]'s proposal, with the preference relation playing the role of the probability distribution, and the reliability relation playing the role of the weight individuals assigns to one another. Also similar is the proposal presented in Lehrer and Wagner [82], where each agent has again both a probability distribution and a weight assignment over each member of the group (including herself), and in which each agent's distribution at a stage $t + 1$ is the linear combination of the distributions at stage t. One remarkable difference is that the latter is more extensively discussed.[21] However, the most important one is that it also studies a setting in which the weights agents assign to one another might change through time, thus allowing a re-evaluation of people by one another at different stages (and even proposing a method for this re-evaluation).

On the other hand, in recent years there have been several proposals studying the way agents influence one another from a logical perspective. Such approaches use qualitative tools instead of quantitative ones, and instead of looking at 'complex' data describing influence behaviour, they rely on relatively 'simple' models, using then formal languages to describe general 'complex' patterns about the system's behaviour. One of such proposals is the study of threshold model dynamics [60] from a logical perspective carried out in Baltag et al. [10], which considers a set of agents, a behaviour p (or a set of them) that agents might or might not have, and a *friendship* relation among the agents themselves. In this setting, each agent will adopt behaviour p if and only if the *proportion* of her p-friends is *at least* the given threshold θ. The proposal of this paper differs from threshold models in that what matters for an agent is not *how many* people has a given preference, but rather the preferences of *certain* people (namely, those she considers the most reliable ones). Still, Baltag et al. [10] goes one step further, as it also incorporates an epistemic component, then

[21] On the conceptual side, Lehrer and Wagner [82] applies its proposal to philosophical problems concerning democracy, justice, science and language. On the technical side, it also explores its axiomatic representation, showing how it satisfies certain conditions of rational consensus.

making an agent's behaviour dependent not on other agents' behaviour, but rather on what she *knows* about other agents' behaviour.

Another related logical approach is the dynamics of peer pressure of Zhen and Seligman [106], whose setting consists on a set of individuals, each one with a preference relation, and a *friendship* relation. The proposal defines two preference-changing operations, one for making a given formula ψ at least as preferable as a given formula φ, $[\varphi \leq \psi]$, and another for making ψ strictly more preferred than φ, $[\varphi < \psi]$. Then it proposes a 'suggestion' dynamics through which an agent will upgrade her preferences with $[\varphi < \psi]$ / $[\varphi \leq \psi]$ if and only if her friends have a strong/weak preference for ψ over φ. Interestingly, the proposal presents an automata describing the way an agent's relative preferences about φ and ψ evolve, and it uses it to characterise those situations in which the iteration of such dynamics will lead to stabilisation (i.e., from some moment on the agents' relative preferences about two given formulas do not change anymore). The main difference between this approach and the proposal of the current paper, besides the reliability relation that defines a hierarchy among an agent's friends, is that while the preference dynamics of the former are about changing the agent's *relative preferences* about two given formulas, the lexicographic upgrade presented here is a 'more semantic' operation that aggregates preference relations. In that sense, this approach is closer to the belief merge of Baltag and Smets [7], which is also about aggregating the full individual preferences (though such approach is rather based on *sequential* communication, with "the order of the speaker" playing the role of this work's reliability relation). Other related logical frameworks are Baltag et al. [9] and Christoff and Hansen [29].

As the reader might have noticed, most of these frameworks rely on a (irreflexive and symmetric) friendship relation, which determines the set of agents that will affect a given agent's future state (her behaviour, her beliefs or her preferences). The general framework presented here does not rely on such relation, as all agents might play a role when determining each agent's new preferences. Nevertheless, such extra component would be useful for the 'non-public' variations explored in Section 4 (in particular, the one of Subsection 4.1), as this friendship relation would provide a semantic (and, by extending the language in the appropriate way, also a syntactic) criteria for characterising the set of agents that will affect the new preferences of a given one. It is also worthwhile to emphasise that none of the frameworks using a friendship relation provides an ordering among such friends, and thus all friends are 'equally important'. Contrastingly, the reliability ordering of this proposal allows to describe not only such situations (by considering all agents equally reliable), but also those in which some agents are more influential than others.[22]

Also related, albeit in a different setting, are the works in the field of belief merging [74; 75; 77], belief combination [12; 13] and judgement aggregation [87; 62], focussing on the aggregation of several independent and potentially conflicting (but equally reliable) sources of information into a consistent one.

[22] Such 'ranked' influence is informally discussed in Liu et al. [89], with the proposed language relying on modalities F_i for 'accessing' friends of a given rank i; this is different from the rather *hybrid* approach [3] used here for formulas describing the reliability relation. Another interesting difference is that, in such proposal, this 'ranking' provides an ordering among friends, thus excluding the agent herself (recall: such friendship relation is asked to be irreflexive). On the other hand, in the setting presented here, each agent should also rank herself.

The main difference between the approaches discussed in previous paragraphs and these ones is that in the latter the information to be aggregated is typically given syntactically, as a set of propositional formulas, different from the semantic approach in which full preference orderings are combined.

6 Conclusions and ongoing/further work

This work studies collective decision making scenarios in which agents share publicly and simultaneously their preferences, and then upgrade them according to both what has been announced and the reliability ordering each one of them assigns to the involved agents. The proposal starts by introducing a framework for representing both the agents' preferences about objects and the relative reliability they assign to one another, together with a rich formal language for describing such structures as well as a sound and complete axiom system characterising its validities. Then, after recalling two priority-based aggregation procedures for the cases in which priority relations are, respectively, total orders and partial orders, it introduces a generalisation for preorder priority relations, then using it in order to upgrade individual preferences, presenting not only a formal language for describing the effects of such upgrades but also a recursion-axiom based axiom system for it that takes advantage of the 'static' language's expressivity.

As it has been discussed, this approach uses ideas from both the *aggregation* and the *deliberative* traditions (the procedure for generating the revised preferences for the former, the acknowledgement that preferences can be influenced by others from the latter). Still, given the limitations it imposes in the process (announcing preferences but not their justifications, and different agents influencing different individuals in different ways), the resulting phenomena is closer to the non-completely rational situations that arise when an individual's thoughts, opinions, feelings, and behaviours are influenced by (the actual, imagined or implied presence of) others. This makes the approach interesting not only for computer science (the combination of conflicting sensors' information received by an agent, the aggregation of multiple databases to build an expert system) but also for economics (strategic interaction in networks) and, of course, social and behavioural sciences (the study of the different ways humans influence one another).

Several questions are still open, some of them specific to the introduced framework. For example, ***(i)*** although there has been a brief exploration to some 'non-public' versions of the lexicographic upgrade, a deeper study of the discussed cases as well as similar ones will allow a better understanding of this procedure in the more general setting in which the agents' information is only partially shared. ***(ii)*** Similarly, an ongoing research work is the study of the long-term effect of the lexicographic upgrade. Can we characterise (syntactically, by extending the formal language; semantically, by describing preference and reliability's initial conditions) those situations in which the repetition of the lexicographic upgrade will lead to full preference unanimity among all agents? What about 'weaker' forms of preference unanimity, as unanimity among a subset of agents about the relative position of a subset of the given options?

Then, the introduced framework can be extended in interesting ways. ***(i)*** So far it only represents the agents' preferences and reliability, but it leaves out epistemic notions. A further extension dealing with the knowledge/beliefs agents have about one another's preferences and reliability would allow the study not only of knowledge-dependent dynamics ("if *I know* the preferences of my most reliable agent, I will upgrade mine accordingly") but also of strategic behaviour ("if *I know* how the rest of the agents will behave, I will behave in a different way") therefore rising the issue of manipulation (e.g., can an agent manipulate the preferences she announces at each stage in order to make her originally most preferred situation also the most preferred one for some set of agents? ***(ii)*** Equally interesting, and as it has been pointed out by some people (and in fact studied in [82] in its probabilistic setting), the announcement of the individual preferences might trigger not only a change in preferences but also a change on reliability. Indeed, after preferences are announced, an agent might actually decide that her formerly most reliable agent should not be the most reliable anymore, either because their preferences are too different, or because there is someone else with 'more similar' preferences (cf. Jirakunkanok et al. [72]). It is worthwhile to explore not only policies and model operations representing such change, but also how they work in tandem with the reliability-based preference dynamics proposed here.

Finally, there is the main idea underlying the framework of this proposal: a combination of deliberation and aggregation for collective decision making. One interesting direction here is to focus not on those particular situations in which the repetition of the presented preference upgrade leads to unanimity, but rather at the properties of the individual preference relations that arise, in general, from such repetitive applications. Some of these properties might allow/facilitate the use of aggregation procedures that suffer from impossibility results in the general case. But, more generally, proposals combining these two strategies will benefit from both, as reasonable combinations of deliberation and aggregation perspectives will provide a more realistic model for group decision making.

Funding The research leading to these results has been partially funded by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement no. 283963.

Appendix

Proof of Proposition 2.1

For the first, the formula $\langle 1\rangle(\varphi \wedge \langle\leq_i\rangle\,\psi)$ indicates the existence of a φ-world from which there is a $\leq_i$-reachable ψ-world, exactly what $\varphi \leq_i^{\exists\exists} \psi$ states. For the second, the formula $[1](\varphi \rightarrow \langle\leq_i\rangle\,\psi)$ indicates that every φ-world in the model can $\leq_i$-reach a ψ-world, exactly what $\varphi \leq_i^{\forall\exists} \psi$ states.

Proof of Proposition 2.2

For the first. **(⇒)** The formula $\langle 1\rangle(\varphi \wedge [>_i]\,\neg\psi)$ indicates the existence of a φ-world w from which all $>_i$-reachable worlds satisfy $\neg\psi$. Let such w be the φ-world required by the $\leq_i^{\exists\forall}$ statement, and consider any ψ-world u. By totality, there are only three possibilities, $w <_i u$, $w \simeq_i u$ or $u <_i w$, but the third cannot be the case: if u is strictly below w, then u would satisfy $\neg\psi$. In the two remaining alternatives, the required $w \leq_i u$ is the case. **(⇐)** The statement indicates the existence of a φ-world w that can $\leq_i$-reach every ψ-world. Let such w be the φ-world required by the formula, and consider any world u that can be $>_i$-reached from w. It cannot satisfy ψ, as otherwise the conditions of the $\leq_i^{\exists\forall}$ statement would be satisfied, so there would be an $\leq_i$ arrow from w to u and thus u would not be reachable from w via $>_i$. Hence, u must satisfy $\neg\psi$.

For the second. **(⇒)** The formula $[1](\psi \to [<_i]\neg\varphi)$ indicates that every ψ-world can $<_i$-reach only $\neg\varphi$-worlds. Take any φ-world w and any ψ-world u. By totality, there are only three possibilities, $w <_i u$, $w \simeq_i u$ or $u <_i w$, but the third cannot be the case: if u is strictly below w, then w would satisfy $\neg\varphi$. In the two remaining alternatives, the required $w \leq_i u$ is the case. **(⇐)** The statement indicates that every φ-world w can $\leq_i$-reach every ψ-world. Take any ψ-world u, and consider any world w that is $<_i$-reachable from u. It cannot satisfy φ, as otherwise the conditions of the $\leq_i^{\forall\forall'}$ statement would be satisfied, so there would be an $\leq_i$-arrow from w to u and thus w would not be reachable from u via $<_i$. Hence, w must satisfy $\neg\varphi$.

Proof of Proposition 2.3

For the first. **(⇒)** The formula $\langle 1\rangle(\varphi \wedge [>_i]\neg\psi)$ indicates the existence of a φ-world w from which all $>_i$-reachable worlds satisfy $\neg\psi$. Let such w be the φ-world required by the $\leq_i^{\exists\forall'}$ statement, and consider any ψ-world u that is $\sim_i$-related to w. There are only three possibilities, $w <_i u$, $w \simeq_i u$ or $u <_i w$, but the third cannot be the case: if u is strictly below w, then u would satisfy $\neg\psi$. In the two remaining alternatives, the required $w \leq_i u$ is the case. **(⇐)** The statement indicates the existence of a φ-world w that can $\leq_i$-reach every ψ-world to which it is $\sim_i$-connected to. Let such w be the φ-world required by the formula, and consider any world u that can be $>_i$-reached from w. It cannot satisfy ψ, as otherwise the conditions of the $\leq_i^{\exists\forall'}$ statement would be satisfied, so there would be an $\leq_i$ arrow from w to u and thus u would not be reachable from w via $>_i$. Hence, u must satisfy $\neg\psi$.

For the second. **(⇒)** The formula $[1](\psi \to [<_i]\neg\varphi)$ indicates that every ψ-world can $<_i$-reach only $\neg\varphi$-worlds. Take any φ-world w and any ψ-world u that is $\sim_i$-related to w. There are only three possibilities, $w <_i u$, $w \simeq_i u$ or $u <_i w$, but the third cannot be the case: if u is strictly below w, then w would satisfy $\neg\varphi$. In the two remaining alternatives, the required $w \leq_i u$ is the case. **(⇐)** The statement indicates that every φ-world w can $\leq_i$-reach every ψ-world to which it is $\sim_i$-connected to. Take any ψ-world u, and consider any world w that is $<_i$-reachable from u. It cannot satisfy φ, as otherwise the conditions of the $\leq_i^{\forall\forall'}$ statement would be satisfied, so there would be an $\leq_i$ arrow from w to u and thus w would not be reachable from u via $<_i$. Hence, w must satisfy $\neg\varphi$.

Proof of Proposition 3.1

Note, first, how the proof makes no assumption about the properties of the preference orderings, and second, how the role played by $\mathtt{A}$'s finiteness can be also represented by $\preccurlyeq$ if such relation is asked to be conversely well-founded.[23]

[23]There is no infinite ascending $\prec$-chain of elements in $\mathtt{A}$.

(⇒) Assume $u \leq_{\preccurlyeq} v$. From Definition 3.1, either ***(1)*** $u \leq_{\mathrm{mn}_{\preccurlyeq}} v$ and, for all $i \in \mathtt{A} \setminus \{\mathrm{mn}_{\preccurlyeq}\}$, $u \simeq_i v$, or else ***(2)*** there is $j \in \mathtt{A} \setminus \{\mathrm{mn}_{\preccurlyeq}\}$ such that $u <_j v$ and, for all $k \in [\prec\rangle(j)$, $u \simeq_k v$. If ***(1)*** is the case, then $u \leq_i v$ holds for all $i \in \mathtt{A}$, and therefore every element of $\mathtt{A}$ satisfies part 1 of Definition 3.2, so $u \leq^{lr}_{\preccurlyeq} v$. Otherwise, ***(2)*** is the case, and hence while all elements of $\mathtt{A} \setminus [\preccurlyeq\rangle(j)$ satisfy part 2 of Definition 3.2 (since $\preccurlyeq$ is a total order, $k \notin [\preccurlyeq\rangle(j)$ implies $k \prec j$, and thus every such k has j, with higher priority, satisfying $u <_j v$), all elements of $[\prec\rangle(j)$ satisfy part 1 (as $u \simeq_k v$ for all $k \in [\prec\rangle(j)$, and thus $u \leq_k v$ for all such k), as it does j itself ($u <_j v$ implies $u \leq_j v$). Since $\preccurlyeq$ is total, $\mathtt{A} = (\mathtt{A} \setminus [\preccurlyeq\rangle(j)) \cup [\prec\rangle(j) \cup \{j\}$ for any j, and then all elements of $\mathtt{A}$ satisfy at least one disjunct in Definition 3.2; therefore, $u \leq^{lr}_{\preccurlyeq} v$.

(⇐) Assume $u \leq^{lr}_{\preccurlyeq} v$. According to Definition 3.2, for all $i \in \mathtt{A}$, ***(1)*** $u \leq_i v$ or ***(2)*** there is $j \in [\prec\rangle(i)$ such that $u <_j v$. Here is a deconstruction of the possible scenarios.

(i) Suppose $u \leq_i v$ for all $i \in \mathtt{A}$. If, additionally, $v \leq_j u$ is the case for all $j \in \mathtt{A} \setminus \{\mathrm{mn}_{\preccurlyeq}\}$ then, $u \simeq_j v$ for all such j. Since $u \leq_{\mathrm{mn}_{\preccurlyeq}} v$, part 1 of Definition 3.1 implies $u \leq_{\preccurlyeq} v$. Otherwise, there is at least one element of $\mathtt{A} \setminus \{\mathrm{mn}_{\preccurlyeq}\}$ who does not consider u at least as plausible as v. Let j be, among such elements, the single one with the highest priority (it exists, as $\mathtt{A}$ is finite and $\preccurlyeq$ is total and antisymmetric); then $v \nleq_j u$ so $u <_j u$ (recall the assumption: $u \leq_i v$ for all $i \in \mathtt{A}$) and, moreover, $v \leq_k u$ for all $k \in [\prec\rangle(j)$. Hence, $u \simeq_k v$ for all such k and therefore, by part 2 of Definition 3.1, $u \leq_{\preccurlyeq} v$.

(ii) Suppose otherwise, i.e., there is at least one element in $\mathtt{A}$ for who does not consider v as at least as plausible as u. Let j be, among such elements, the one with the highest priority, so $u \nleq_j v$ and, for all $k \in [\prec\rangle(j)$, $u \leq_k v$. From Definition 3.2, there should be at least one element of $\mathtt{A}$ with higher priority than j that prefers v over u; let j' be, among such elements, again the one with the highest priority, so $j \prec j'$, $u <_{j'} v$ and, for all $k \in [\prec\rangle(j')$, $u \nless_k v$. Consider now the elements of $[\prec\rangle(j')$: for each such k, $u \nless_k v$ and thus either $u \nleq_k v$ or else $v \leq_k u$. But the first disjunct fails for all of them: $k \in [\prec\rangle(j')$ implies $j' \prec k$ and then, since $j \prec j'$, transitivity implies $j \prec k$ so $k \in [\prec\rangle(j)$ and hence $u \leq_k v$. Then, for all such k, both $u \leq_k v$ (as they have higher priority than j) and $v \leq_k u$ (as they have higher priority than j'), so $u \simeq_k v$. Thus, there is j' in $\mathtt{A} \setminus \{\mathrm{mn}_{\preccurlyeq}\}$ (as j' is strictly above j) such that $u <_{j'} v$ and every $k \in [\prec\rangle(j')$ is such that $u \simeq_k v$; hence, by part 2 of Definition 3.1, $u \leq_{\preccurlyeq} v$.

Proof of Proposition 3.2

(⇒) Assume $u \leq^{lr}_{\preccurlyeq} v$. Then, from Definition 3.2, for all $i \in \mathtt{A}$, $u \leq_i v$ or else there is $j \in [\prec\rangle(i)$ such that $u <_j v$. But in the first case, i satisfies part 1 of Definition 3.3. And in the second, the set $\{j\} \subseteq [\prec\rangle(i)$ is such that ***(2.1)*** $j \approxeq j$, as $\preccurlyeq$ is reflexive, ***(2.2)*** for all $k \in \mathtt{A} \setminus \{j\}$, $j \not\approxeq k$, as $\preccurlyeq$ is antisymmetric, and ***(2.3)*** $u <_j v$, so i satisfies part 2 of Definition 3.3. Hence, $u \leq_{\preccurlyeq} v$.

(⇐) Assume $u \leq_{\preccurlyeq} v$. Then, from Definition 3.3, for all $i \in \mathtt{A}$, $u \leq_i v$ or else there is $\mathtt{B}$ such that $\mathtt{B} \subseteq [\prec\rangle(i)$ and ***(2.1)*** $j_1 \preccurlyeq j_2$ for all $j_1, j_2 \in \mathtt{B}$, ***(2.2)*** for every $k \in \mathtt{A} \setminus \mathtt{B}$ there is a $j \in \mathtt{B}$ such that $j \not\approxeq k$, and ***(2.3)*** $u <_j v$ for all $j \in \mathtt{B}$. In the first case, i satisfies part 1 of Definition 3.2. In the second, note how any such set $\mathtt{B}$ has at most one element, as there cannot be different $j_1, j_2 \in \mathtt{B}$ satisfying $j_1 \preccurlyeq j_2$

because $\preccurlyeq$ is antisymmetric, and it is non-empty, as for every element not in $\mathtt{B}$ (and there is at least one, since $\mathtt{A}$ contains at least two different agents) there should be an element in $\mathtt{B}$ satisfying the given condition. Thus, $\mathtt{B}$ is a singleton, and then its single element j is such that $j \in [\prec\rangle(i)$ and $u <_j v$. Thus, i satisfies part 2 of part 1 of Definition 3.2. Hence, $u \leq^{lr}_{\preccurlyeq} v$.

Proof of Proposition 3.3

First, a useful lemma: if $u \nleq_i v$ for all $i \in \mathtt{A}$, then $u \nleq_{\preccurlyeq} v$. For its proof, take any $i \in \mathtt{A}$. The first disjunct in Definition 3.3 clearly fails, as $u \nleq_i v$. But the second disjunct asks for a non-empty cluster $\mathtt{B} \subseteq \mathtt{A}$ whose elements put v above u, and this cannot be the case as $u \nleq_i v$ for all $i \in \mathtt{A}$, and thus $u \nless_i v$ for all such agents. Hence, as the disjunction fails for every agent in (the non-empty) $\mathtt{A}$, $u \nleq_{\preccurlyeq} v$.

Now the proof of the proposition. ***(i)*** If $u \simeq_i v$ for all $i \in \mathtt{A}$, then both $u \leq_i v$ and $v \leq_i u$ for all such agents; hence, $u \leq_{\preccurlyeq} v$ and $v \leq_{\preccurlyeq} u$, and therefore $u \simeq_{\preccurlyeq} v$. ***(ii)*** If $u <_i v$ for all $i \in \mathtt{A}$, then both $u \leq_i v$ and $v \nleq_i u$ for all such agents. From the first, $u \leq_{\preccurlyeq} v$; from the second and the previous lemma, $v \nleq_{\preccurlyeq} u$. Therefore, $u <_{\preccurlyeq} v$. ***(iii)*** If $u \nsim_i v$ for all $i \in \mathtt{A}$, then both $u \nleq_i v$ and $v \nleq_i u$ for all such agents; hence, from the previous lemma, $u \nleq_{\preccurlyeq} v$ and $v \nleq_{\preccurlyeq} u$, and therefore $u \nsim_{\preccurlyeq} v$.

Proof of Proposition 3.4

For reflexivity, take any $u \in W$. Since $\leq_i$ is reflexive for every $i \in \mathtt{A}$, then $u \leq_i u$ for all such i, so every element of $\mathtt{A}$ satisfy the first disjunct of the definition; then, $u \leq_{\preccurlyeq} u$.

For transitivity[24], suppose $u \leq_{\preccurlyeq} v$ and $v \leq_{\preccurlyeq} w$. From the first, for every $i \in \mathtt{A}$, either $(\mathbf{1}_{u,v})$ $u \leq_i v$ or else $(\mathbf{2}_{u,v})$ there is a non-empty $\preccurlyeq$-cluster $\prec$-above i whose elements prefer v over u; from the second, for every $i \in \mathtt{A}$, either $(\mathbf{1}_{v,w})$ $v \leq_i w$ or else $(\mathbf{2}_{v,w})$ there is a non-empty $\preccurlyeq$-cluster $\prec$-above i whose elements prefer w over v. Now, take any $i \in \mathtt{A}$.

- ***(i)*** Suppose i satisfies $(\mathbf{1}_{u,v})$ and $(\mathbf{1}_{v,w})$. Then $u \leq_i v$ and $v \leq_i w$ and thus, by transitivity, $u \leq_i w$, that is, i satisfies $(\mathbf{1}_{u,w})$ and therefore it satisfies the disjunction in Definition 3.3 for the pair u, w.
- ***(ii)*** Suppose i satisfies $(\mathbf{2}_{u,v})$, i.e., suppose there is a non-empty $\preccurlyeq$-cluster $\prec$-above i whose elements prefer v over u. Let $\mathtt{B}_1$ be, among all such clusters, a $\preccurlyeq$-maximal one (which exists by $\mathtt{A}$'s finiteness; thus, not only $j \in \mathtt{B}_1$ implies $u <_j v$, but also every cluster $\prec$-above $\mathtt{B}_1$ contains at least one element that does not prefer v over u), and consider its elements' opinion for the pair v, w.
  - ***(a)*** If $v \leq_j w$ for every $j \in \mathtt{B}_1$ then, as $u <_j v$ for all such j, it follows from transitivity that $u \leq_j w$ for each j. Moreover, $w \nleq_j u$ for each such j, as otherwise $w \leq_j u$, $v \leq_j w$ and transitivity would imply $v \leq_j u$, contradicting $u <_j v$. Then $u <_j w$ for each $j \in \mathtt{B}_1$ and hence, as $\mathtt{B}_1$ is a non-empty $\preccurlyeq$-cluster $\prec$-above i whose elements prefer w over u, i satisfies $(\mathbf{2}_{u,w})$ and therefore it satisfies the disjunction in Definition 3.3 for the pair u, w.

[24] Cf. Grosof [61].

(b) Otherwise, there is an element of $\mathtt{B}_1$ for whom w is not at least as preferable as v, $v \not\leq w$. But $v \leq_{\preccurlyeq} w$ is the case, and therefore there should be a non-empty $\preccurlyeq$-cluster $\prec$-above such element (thus above the cluster $\mathtt{B}_1$ and hence above the initial i) whose elements prefer w over v. Let $\mathtt{B}_2$ be one of such clusters.

Now, recall that $\mathtt{B}_1$ is a $\preccurlyeq$-maximal cluster with all its elements preferring v over u. Then, since every cluster $\prec$-above $\mathtt{B}_2$ is also $\prec$-above $\mathtt{B}_1$, every cluster $\prec$-above $\mathtt{B}_2$ has at least one element for which $u \not< v$; therefore, every $j \in \mathtt{B}_2$ should satisfy $u \leq_j v$, as otherwise any failing j would also fail to satisfy the disjunction of Definition 3.3 for the pair u, v, and hence the very initial assumption $u \leq_{\preccurlyeq} v$ would fail. Thus, $j \in \mathtt{B}_2$ implies not only $v <_j w$ but also $u \leq_j v$ and hence, by transitivity, $u \leq_j w$. Finally, observe how $w \not\leq_j u$ should be the case for every such j, as otherwise $w \leq_j u$, $u \leq_j v$ and transitivity would imply $w \leq_j v$, contradicting $v <_j w$; then, $u <_j w$ for all $j \in \mathtt{B}_2$. Thus, as $\mathtt{B}_2$ is a non-empty $\preccurlyeq$-cluster $\prec$-above i whose elements prefer w over u, i satisfies **($2_{u,w}$)** and therefore it satisfies the disjunction in Definition 3.3 for the pair u, w.

(iii) The case in which i satisfies **($2_{v,w}$)** is analogous to the previous one.

Thus, every $i \in \mathtt{A}$ satisfies the disjunction in Definition 3.3 for the pair u, w. Therefore, $u \leq_{\preccurlyeq} w$.

Proof of Proposition 3.5

Suppose, for the sake of contradiction, that there are $u, v \in W$ with $u \neq v$ such that $u \leq_{\preccurlyeq} v$ and $v \leq_{\preccurlyeq} u$. From the first, for every $i \in \mathtt{A}$, **($1_{u,v}$)** $u \leq_i v$ or else **($2_{u,v}$)** there is a non-empty $\preccurlyeq$-cluster $\prec$-above i whose elements prefer v over u; from the second, for every $i \in \mathtt{A}$, either **($1_{v,u}$)** $v \leq_i u$ or else **($2_{v,u}$)** there is a non-empty $\preccurlyeq$-cluster $\prec$-above i whose elements prefer u over v. Now, take any such $i \in \mathtt{A}$ (which exists, as $\mathtt{A}$ is non-empty). Since $\leq_i$ is antisymmetric and u, v are different, i cannot satisfy both **($1_{u,v}$)** and **($1_{v,u}$)**; hence, it should satisfy **($2_{u,v}$)** or **($2_{v,u}$)**.

If i satisfies **($2_{u,v}$)**, then there is a non-empty $\preccurlyeq$-cluster $\prec$-above i whose elements prefer v over u; let $\mathtt{B}_1$ be, among all such clusters, a $\preccurlyeq$-maximal one (so not only $j \in \mathtt{B}_1$ implies $u <_j v$, but also every cluster $\prec$-above $\mathtt{B}_1$ contains at least one element that does not prefer v over u), and take any of its elements, say j_1. Since $j_1 \in \mathtt{B}_1$, then $u <_{j_1} v$, that is, $v \not\leq_{j_1} u$; but $v \leq_{\preccurlyeq} u$ is the case, and thus there should be a non-empty $\preccurlyeq$-cluster $\prec$-above j_1 (and thus $\prec$-above $\mathtt{B}_1$) whose elements prefer u over v. Let $\mathtt{B}_2$ be, among all such clusters, a $\preccurlyeq$-maximal one (so not only $j \in \mathtt{B}_2$ implies $v <_j u$, but also every cluster $\prec$-above $\mathtt{B}_2$ contains at least one element that does not prefer u over v), and take any of its elements, say j_2. Since $j_2 \in \mathtt{B}_2$, then $v <_{j_2} u$, that is, $u \not\leq_{j_1} v$; then, as $u \leq_{\preccurlyeq} v$, there should be a non-empty $\preccurlyeq$-cluster $\prec$-above j_2 (and thus $\prec$-above $\mathtt{B}_1$) whose elements prefer v over u. But this cannot be the case, as any such cluster, being also $\prec$-above $\mathtt{B}_1$, should have at least at least one element that does not prefer v over u. Thus, i cannot satisfy **($2_{u,v}$)**. A similar argument shows that i cannot satisfy **($2_{v,u}$)** either. Thus, the initial assumption fails: there are no $u, v \in W$ with $u \neq v$ such that $u \leq_{\preccurlyeq} v$ and $v \leq_{\preccurlyeq} u$.

Proof of Proposition 3.6

From left to right, suppose $u \leq_{\preccurlyeq} v$. If $u \leq_i v$ holds for all $i \in \mathtt{A}$, then so does part 1 of the right-hand side, and thus the whole right-hand side holds. Otherwise, from $\leq_{\preccurlyeq}$'s definition (Definition 3.3), for each i with $u \nleq_i v$ there is a non-empty $\preccurlyeq$-cluster lying $\prec$-above i whose agents put all v above u. Since $\preccurlyeq$ is total and $\mathtt{A}$ is finite, among all such clusters it is possible to select the one with higher priority, $\mathtt{B}$, so no cluster above $\mathtt{B}$ contains only elements placing v above u. Clearly, $j \in \mathtt{B}$ implies $u <_j v$; moreover, there is no $j' \in \mathtt{B}'$ with $\mathtt{B}' \in C^{\mathtt{B}}_{\preccurlyeq}$ such that $u \nleq_{j'} v$ as otherwise, since no cluster above $\mathtt{B}$ contains only elements placing v above u, j' would fail in $\leq_{\preccurlyeq}$'s definition, and thus $u \nleq_{\preccurlyeq} v$. Hence, part 2 of the right-hand side holds, and so does whole right-hand side.

From right to left, suppose part 1 of the right-hand side holds; then, clearly, $u \leq_{\preccurlyeq} v$. Otherwise, part 2 of the right-hand side holds, i.e., there is a non-empty cluster of agents $\mathtt{B}$ whose elements place v above u, with every element of every cluster with higher priority considering v at least as preferable as u. Now consider any agent $i \in \mathtt{A}$: if it is in $\mathtt{B}$ or in a cluster above $\mathtt{B}$, then $u \leq_i v$; otherwise, it is in a cluster below $\mathtt{B}$, but then $\mathtt{B}$ itself is the needed non-empty $\preccurlyeq$-cluster lying $\prec$-above i whose agents put all v above u. Thus, all agents in $\mathtt{A}$ satisfy the disjunction in Definition 3.3 and therefore $u \leq_{\preccurlyeq} v$.

Proof of Theorem 2

Case 1. By $\langle\!\langle \cdot \rangle\!\rangle$'s and M_{lx}'s definitions, $\langle\!\langle 1 \rangle\!\rangle^{M_{\text{lx}}} = W \times W = \langle\!\langle 1 \rangle\!\rangle^{M} = \langle\!\langle Tx(1) \rangle\!\rangle^{M}$.

Case $\geq_a$. Analogous to that for $\leq_a$.

Case $?(\varphi, \psi)$. It follows from the definitions:

$$\begin{aligned}\langle\!\langle ?(\varphi,\psi) \rangle\!\rangle^{M_{\text{lx}}} &= [\![\varphi]\!]^{M_{\text{lx}}} \times [\![\psi]\!]^{M_{\text{lx}}} = [\![\langle \text{lx} \rangle\, \varphi]\!]^{M} \times [\![\langle \text{lx} \rangle\, \psi]\!]^{M} = \langle\!\langle ?\,(\langle \text{lx} \rangle\, \varphi, \langle \text{lx} \rangle\, \psi) \rangle\!\rangle^{M} \\ &= \langle\!\langle Tx(?(\varphi,\psi)) \rangle\!\rangle^{M}\end{aligned}$$

Case $-\pi$. It follows from the definitions and the inductive hypothesis $\langle\!\langle \pi \rangle\!\rangle^{M_{\text{lx}}} = \langle\!\langle Tx(\pi) \rangle\!\rangle^{M}$:

$$\begin{aligned}\langle\!\langle -\pi \rangle\!\rangle^{M_{\text{lx}}} &= (W \times W) \setminus \langle\!\langle \pi \rangle\!\rangle^{M_{\text{lx}}} = (W \times W) \setminus \langle\!\langle Tx(\pi) \rangle\!\rangle^{M} = \langle\!\langle -Tx(\pi) \rangle\!\rangle^{M} \\ &= \langle\!\langle Tx(-\pi) \rangle\!\rangle^{M}\end{aligned}$$

Case $\pi \cup \sigma$. It follows from the definitions and the inductive hypotheses $\langle\!\langle \pi \rangle\!\rangle^{M_{\text{lx}}} = \langle\!\langle Tx(\pi) \rangle\!\rangle^{M}$ and $\langle\!\langle \sigma \rangle\!\rangle^{M_{\text{lx}}} = \langle\!\langle Tx(\sigma) \rangle\!\rangle^{M}$:

$$\begin{aligned}\langle\!\langle \pi \cup \sigma \rangle\!\rangle^{M_{\text{lx}}} &= \langle\!\langle \pi \rangle\!\rangle^{M_{\text{lx}}} \cup \langle\!\langle \sigma \rangle\!\rangle^{M_{\text{lx}}} = \langle\!\langle Tx(\pi) \rangle\!\rangle^{M} \cup \langle\!\langle Tx(\sigma) \rangle\!\rangle^{M} = \langle\!\langle Tx(\pi) \cup Tx(\sigma) \rangle\!\rangle^{M} \\ &= \langle\!\langle Tx(\pi \cup \sigma) \rangle\!\rangle^{M}\end{aligned}$$

Case $\pi \cap \sigma$. Analogous to the previous one.

References

[1] M. Altman, editor. *Real-World Decision Making. An Encyclopedia of Behavioral Economics*. Greenwood, California, USA, 2015. ISBN 978-1-4408-2815-7. Cited on page 2.

[2] H. Andréka, M. D. Ryan, and P.-Y. Schobbens. Operators and laws for combining preference relations. *Journal of Logic and Computation*, 12(1):13–53, 2002. DOI: 10.1093/logcom/12.1.13. URL: `ftp://ftp.cs.bham.ac.uk/pub/authors/M.D.Ryan/97-operators.ps`. Cited on pages **5**, **15**, **and 19**.

[3] C. Areces and B. ten Cate. Hybrid logics. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*, volume 3 of *Studies in Logic and Practical Reasoning*, pages 821–868. Elsevier Science Inc., Amsterdam, 2006. Cited on page **33**.

[4] K. J. Arrow. *Social Choice and Individual Values*. John Wiley & Sons, Inc., 2nd edition, 1963. First edition published in 1951. Cited on page **1**.

[5] K. J. Arrow, A. K. Sen, and K. Suzumura, editors. *Handbook of Social Choice and Welfare*, volume 1. Elsevier, Aug. 2002. ISBN 978-0-444-82914-6. Cited on pages **1 and 7**.

[6] K. J. Arrow, A. K. Sen, and K. Suzumura, editors. *Handbook of Social Choice and Welfare*, volume 2. Elsevier, Sept. 2010. ISBN 978-0-444-50894-2. Cited on pages **1 and 7**.

[7] A. Baltag and S. Smets. Protocols for belief merge: Reaching agreement via communication. *Logic Journal of the IGPL*, 21(3):468–487, 2013. DOI: 10.1093/jigpal/jzs049. Cited on page **33**.

[8] A. Baltag, L. S. Moss, and S. Solecki. The logic of public announcements, common knowledge, and private suspicions. In I. Gilboa, editor, *TARK*, pages 43–56, San Francisco, CA, USA, 1998. Morgan Kaufmann. ISBN 1-55860-563-0. URL: `http://dl.acm.org/citation.cfm?id=645876.671885`. Cited on page **25**.

[9] A. Baltag, Z. Christoff, J. U. Hansen, and S. Smets. Logical models of informational cascades. In J. van Benthem and F. Liu, editors, *Logic Across the University: Foundations and Applications. Proceedings of the Tsinghua Logic Conference, Beijing, 2013*, volume 47 of *Studies in Logic*, pages 405–432, London, 2013. College Publications. ISBN 978-1-84890-122-3. URL: `https://zoechristoff.files.wordpress.com/2012/09/logical-models-of-information-cascades.pdf`. Cited on page **33**.

[10] A. Baltag, Z. Christoff, R. K. Rendsvig, and S. Smets. Dynamic epistemic logics of diffusion and prediction in social networks. Technical Report PP-2015-22, Institute for Logic, Language and Computation, University of Amsterdam, 2015. URL: `https://www.illc.uva.nl/Research/Publications/Reports/PP-2015-22.text.pdf`. Cited on page **32**.

[11] A. V. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817, 1992. DOI: 10.2307/2118364. URL: `http://economics.mit.edu/files/8869`. Cited on page **2**.

[12] C. Baral, S. Kraus, and J. Minker. Combining multiple knowledge bases. *IEEE Transactions on Knowledge and Data Engineering*, 3(2):208–220, 1991. DOI: 10.1109/69.88001. Cited on page **33**.

[13] C. Baral, S. Kraus, J. Minker, and V. S. Subrahmanian. Combining knowledge bases consisting of first-order analysis. *Computational Intelligence*, 8:45–71, 1992. DOI: 10.1111/j.1467-8640.1992.tb00337.x. Cited on page **33**.

[14] S. Barberà, W. Bossert, and P. K. Pattanaik. Ranking sets of objects. In S. Barberà, P. J. Hammond, and C. Seidl, editors, *Handbook of Utility Theory*, volume 2: Extensions, pages 893–977. Kluwer Academic Publisher, Boston, 2004. ISBN 1-4020-7714-9. DOI: 10.1007/978-1-4020-7964-1_4. Cited on page 11.

[15] J. van Benthem. Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics*, 17(2):129–155, 2007. DOI: 10.3166/jancl.17.129-155. URL: `http://www.illc.uva.nl/Publications/ResearchReports/PP-2006-11.text.pdf`. Cited on page 21.

[16] J. van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, 2011. ISBN 978-0-521-76579-4. Cited on page 24.

[17] J. van Benthem, J. van Eijck, and B. Kooi. Logics of communication and change. *Information and Computation*, 204(11):1620–1662, Nov. 2006. ISSN 0890-5401. DOI: 10.1016/j.ic.2006.04.006. URL: `http://www.illc.uva.nl/Publications/ResearchReports/PP-2005-09.text.pdf`. Cited on page 25.

[18] J. van Benthem, S. van Otterloo, and O. Roy. Preference logic, conditionals and solution concepts in games. In Lagerlund et al. [79], pages 61–76. URL: `http://www.illc.uva.nl/Publications/ResearchReports/PP-2005-28.text.pdf`. Cited on page 11.

[19] J. van Benthem, P. Girard, and O. Roy. Everything else being equal: A modal logic for ceteris paribus preferences. *Journal of Philosophical Logic*, 38(1):83–125, 2009. ISSN 0022-3611. DOI: 10.1007/s10992-008-9085-3. URL: `http://www.illc.uva.nl/Publications/ResearchReports/PP-2007-09.text.pdf`. Cited on page 11.

[20] J. F. A. K. van Benthem. Minimal deontic logic (abstract). *Bulletin of the Section of Logic*, 8(1):36–41, 1979. Cited on page 10.

[21] S. Bikhchandani, D. Hirshleifer, and I. Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100 (5):992–1026, 1992. DOI: 10.1086/261849. URL: `http://www.jstor.org/stable/2138632`. Cited on page 2.

[22] L. Blume, A. Brandenburger, and E. Dekel. Lexicographic probabilities and choice under uncertainty. *Econometrica*, 59(1):61–79, Jan. 1991. DOI: 10.2307/2938240. URL: `http://www.jstor.org/stable/2938240`. Cited on page 4.

[23] J. Bohman. The coming of age of deliberative democracy. *The Journal of Political Philosophy*, 6(4):400–425, Dec. 1998. ISSN 1467-9760. DOI: 10.1111/1467-9760.00061. Cited on page 2.

[24] A. Brandenburger, A. Friedenberg, and H. J. Keisler. Admissibility in games. *Journal of Economic Theory*, 76(2):307–352, Mar. 2008. DOI: 10.1111/j.1468-0262.2008.00835.x. Cited on page 4.

[25] J. P. Burgess. Basic tense logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume II, chapter 2, pages 89–133. Reidel, Dordrecht, 1984. Cited on pages 9 **and** 13.

[26] J. Cantwell. Resolving conflicting information. *Journal of Logic, Language and Information*, 7(2):191–220, 1998. DOI: 10.1023/A:1008296216319. Cited on page 8.

[27] C. Castelfranchi and R. Falcone. Principles of trust for MAS: Cognitive anatomy, social importance, and quantification. In Y. Demazeau, editor, *ICMAS*, pages 72–79. IEEE Computer Society, 1998. ISBN 0-8186-8500-X. DOI: 10.1109/ICMAS.1998.699034. Cited on page 8.

[28] C. Castelfranchi and Y.-H. Tan, editors. *Trust and Deception in Virtual Societies*. Springer Netherlands, Dordrecht, 2001. ISBN 978-94-017-3614-5. DOI: 10.1007/978-94-017-3614-5. Cited on pages **43 and 44**.

[29] Z. Christoff and J. U. Hansen. A logic for diffusion in social networks. *Journal of Applied Logic*, 13(1):48–77, 2015. DOI: 10.1016/j.jal.2014.11.011. Cited on page **33**.

[30] J. Cohen. Deliberation and democratic legitimacy. In D. Matravers and J. E. Pike, editors, *Debates in contemporary political philosophy: An anthology*, pages 342–360. Routledge, London, 2003. ISBN 0415302102. Reprinted from [68], pages 17–34. Cited on page **2**.

[31] M. H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974. URL: `http://www.jstor.org/stable/2285509`. Cited on pages **3 and 32**.

[32] E. Dekel, A. Friedenberg, and M. Siniscalchi. Lexicographic beliefs and assumption. *Journal of Economic Theory*, 163(3):955–985, May 2016. DOI: 10.1016/j.jet.2016.03.003. Cited on page **4**.

[33] R. Demolombe. To trust information sources: A proposal for a modal logic framework. In Castelfranchi and Tan [28], pages 111–124. ISBN 978-94-017-3614-5. DOI: 10.1007/978-94-017-3614-5_5. Cited on page **8**.

[34] R. Demolombe. Reasoning about trust: A formal logical framework. In C. D. Jensen, S. Poslad, and T. Dimitrakos, editors, *iTrust*, volume 2995 of *Lecture Notes in Computer Science*, pages 291–303. Springer, 2004. ISBN 3-540-21312-0. DOI: 10.1007/978-3-540-24747-0_22. Cited on page **8**.

[35] R. Demolombe and C.-J. Liau. A logic of graded trust and belief fusion. In *Proceedings of the 4th Workshop on Deception, Fraud and Trust in Agent Societies*, pages 13–25, 2001. Cited on page **8**.

[36] H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*, volume 337 of *Synthese Library Series*. Springer, Dordrecht, The Netherlands, 2008. ISBN 978-1-4020-5838-7. Cited on page **24**.

[37] J. S. Dryzek. *Deliberative Democracy and Beyond: Liberals, Critics, Contestations*. Oxford University Press, New York, 2000. ISBN 0-19-829507-3. Cited on page **2**.

[38] J. S. Dryzek and C. List. Social choice theory and deliberative democracy: A reconciliation. *British Journal of Political Science*, 33(1):1–28, Jan. 2003. ISSN 0007-1234. DOI: 10.1017/S0007123403000012. Cited on page **2**.

[39] K. Durkin. Peer pressure. In A. S. R. Manstead and M. Hewstone, editors, *The Blackwell Encyclopedia of Social Psychology*. Blackwell, 1995. ISBN 978-0-631-20289-9. Cited on page **2**.

[40] D. Easley and J. Kleinberg. *Networks, Crowds and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, New York, 2010. ISBN 978-0-521-19533-1. Cited on page **2**.

[41] J. Elster. The market and the forum: three varieties of political theory. In J. Elster and A. Hylland, editors, *Foundations of Social Choice Theory*. Cambridge University Press, Cambridge, 1986. Cited on page **2**.

[42] U. Endriss. Logic and social choice theory. In A. Gupta and J. van Benthem, editors, *Logic and Philosophy Today*, volume 2, pages 333–377. College Publications, 2011. ISBN 978-1-84890-041-7. Cited on page **1**.

[43] R. Falcone and C. Castelfranchi. Social trust: A cognitive approach. In Castelfranchi and Tan [28], pages 55–90. ISBN 978-94-017-3614-5. DOI: 10.1007/978-94-017-3614-5_3. Cited on page 8.

[44] R. Falcone, M. P. Singh, and Y.-H. Tan, editors. *Trust in Cyber-societies, Integrating the Human and Artificial Perspectives [based on a workshop on Deception, Fraud, and Trust in Agent Societies held during the Autonomous Agents Conference in Barcelona, Spain in June 2000]*, volume 2246 of *Lecture Notes in Computer Science*, 2001. Springer. ISBN 3-540-43069-5. DOI: 10.1007/3-540-45547-7. Cited on page 8.

[45] R. Falcone, K. S. Barber, J. Sabater-Mir, and M. P. Singh, editors. *Trust in Agent Societies, 11th International Workshop, TRUST 2008, Estoril, Portugal, May 12-13, 2008. Revised Selected and Invited Papers*, volume 5396 of *Lecture Notes in Computer Science*, 2008. Springer. ISBN 978-3-540-92802-7. DOI: 10.1007/978-3-540-92803-4. Cited on pages **8 and 47**.

[46] P. C. Fishburn. A study of lexicographic expected utility. *Management Science*, 17 (11):672–678, July 1971. DOI: 10.1287/mnsc.17.11.672. URL: `http://www.jstor.org/stable/2629309`. Cited on page 4.

[47] P. C. Fishburn. Lexicographic orders, utilities and decision rules: A survey. *Management Science*, 20(11):1442–1471, July 1974. DOI: 10.1287/mnsc.20.11.1442. URL: `http://www.jstor.org/stable/2629975`. Cited on page 4.

[48] J. R. P. French. A formal theory of social power. *Psychological Review*, 63(3):181–194, 1956. DOI: 10.1037/h0046123. Cited on page 32.

[49] D. M. Gabbay, O. Rodrigues, and G. Pigozzi. Connections between belief revision, belief merging and social choice. *Journal of Logic and Computation*, 19(3):445–446, 2009. DOI: 10.1093/logcom/exn013. Cited on page 2.

[50] G. Gargov and S. Passy. A note on boolean modal logic. In P. P. Petkov, editor, *Mathematical Logic*, pages 311–321. Plenum Press, 1990. DOI: 10.1007/978-1-4613-0609-2_21. Cited on pages **9 and 13**.

[51] G. Gargov, S. Passy, and T. Tinchev. Modal environment for boolean speculations. In D. Skordev, editor, *Mathematical Logic and Its Applications*, pages 253–263. Plenum Press, 1987. DOI: 10.1007/978-1-4613-0897-3_17. Cited on page 10.

[52] S. Ghosh and F. R. Velázquez-Quesada. A note on reliability-based preference dynamics. In W. van der Hoek, W. H. Holliday, and W. Wang, editors, *Logic, Rationality, and Interaction - 5th International Workshop, LORI 2015 Taipei, Taiwan, October 28-31, 2015, Proceedings*, volume 9394 of *Lecture Notes in Computer Science*, pages 129–142. Springer, 2015. ISBN 978-3-662-48560-6. DOI: 10.1007/978-3-662-48561-3_11. URL: `http://bit.ly/1Jv30y3`. Cited on pages **4, 9, 13, and 21**.

[53] S. Ghosh and F. R. Velázquez-Quesada. Agreeing to agree: Reaching unanimity via preference dynamics based on reliable agents. In G. Weiss, P. Yolum, R. H. Bordini, and E. Elkind, editors, *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2015, Istanbul, Turkey, May 4-8, 2015*, pages 1491–1499. ACM, 2015. ISBN 978-1-4503-3413-6. URL: `http://dl.acm.org/citation.cfm?id=2773342`. URL: `http://bit.ly/1Fq49Vm`. Cited on pages **4, 5, 14, 21, and 26**.

[54] P. Girard. *Modal Logic for Belief and Preference Change*. PhD thesis, Department of Philosophy, Stanford University, Stanford, CA, USA, Feb. 2008. ILLC Dissertation Series DS-2008-04. Cited on page 11.

[55] J. Golbeck, editor. *Computing with Social Trust*. Human-Computer Interaction Series. Springer, 2009. ISBN 978-1-84800-355-2. DOI: 10.1007/978-1-84800-356-9. Cited on page **8**.

[56] R. Goldbach. Modelling democratic deliberation. Master's thesis, Institute for Logic, Language and Computation (ILLC), Universiteit van Amsterdam (UvA), Amsterdam, The Netherlands, 2015. URL: `http://www.illc.uva.nl/Research/Publications/Reports/MoL-2015-05.text.pdf`. ILLC Master of Logic Thesis Series MoL-2015-05. Cited on page **2**.

[57] R. Goldblatt. *Logics of Time and Computation*. Number 7 in CSLI Lecture Notes. CSLI Publications, Stanford, CA, 2nd edition, 1992. ISBN 978-0-937073-94-0. Cited on page **9**.

[58] R. I. Goldblatt. Semantic analysis of orthologic. *Journal of Philosophical Logic*, 3 (1-2):19–35, Mar. 1974. ISSN 0022-3611. DOI: 10.1007/BF00652069. Cited on page **10**.

[59] A. I. Goldman. Experts: Which ones should you trust? *Philosophy and Phenomenological Research*, 63(1):85–110, 2001. ISSN 1933-1592. DOI: 10.1111/j.1933-1592.2001.tb00093.x. Cited on page **8**.

[60] M. Granovetter. Threshold models of collective behavior. *American Journal of Sociology*, 83(6):1420–1443, May 1978. DOI: 10.1086/226707. URL: `http://www.jstor.org/stable/2778111`. Cited on page **32**.

[61] B. N. Grosof. Generalizing prioritization. In J. F. Allen, R. Fikes, and E. Sandewall, editors, *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR'91). Cambridge, MA, USA, April 22-25, 1991.*, pages 289–300, Cambridge, MA, USA, Apr. 1991. Morgan Kaufmann. ISBN 1-55860-165-1. Cited on pages **4 and 38**.

[62] D. Grossi and G. Pigozzi. *Judgment Aggregation: A Primer*. Number 27 in Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2014. ISBN 9781627050876. DOI: 10.2200/S00559ED1V01Y201312AIM027. Cited on page **33**.

[63] T. Grüne-Yanoff and S. O. Hansson. Preference change: An introduction. In *Preference Change* Grüne-Yanoff and Hansson [64], pages 1–26. ISBN 978-90-481-2593-7. DOI: 10.1007/978-90-481-2593-7_1. Cited on page **5**.

[64] T. Grüne-Yanoff and S. O. Hansson, editors. *Preference Change*, volume 42 of *Theory and Decision Library*. Springer, 2009. ISBN 978-90-481-2593-7. DOI: 10.1007/978-90-481-2593-7. Cited on pages **2, 7, and 45**.

[65] A. Gutmann and D. Thompson. *Democracy and Disagreement*. Belknap Press of Harvard University Press, Cambridge, Mass., 1996. ISBN 0-674-19766-6. Cited on page **2**.

[66] A. Gutmann and D. Thompson. *Why Deliberative Democracy?* Princeton University Press, Princeton, New Jersey, 2004. ISBN 0-691-12019-6. Cited on page **2**.

[67] J. Habermas. *Between Facts and Norms: Contributions to a Discursive Theory of Law and Democracy*. MIT Press, Cambridge, 1996. Cited on page **2**.

[68] A. P. Hamlin and P. Pettit, editors. *The Good polity: Normative analysis of the state*. Basil Blackwell, Oxford, 1989. ISBN 0631158049. Cited on page **43**.

[69] D. Harel, D. Kozen, and J. Tiuryn. *Dynamic Logic*. MIT Press, Cambridge, MA, 2000. ISBN 0-262-08289-6. Cited on pages **10 and 25**.

[70] A. Herzig, E. Lorini, J. F. Hübner, and L. Vercouter. A logic of trust and reputation. *Logic Journal of the IGPL*, 18(1):214–244, 2010. DOI: 10.1093/jigpal/jzp077. Cited on page **8**.

[71] W. H. Holliday. Trust and the dynamics of testimony. In D. Grossi, L. Kurzen, and F. R. Velázquez-Quesada, editors, *Logic and Interactive RAtionality. Seminar's yearbook 2009*, pages 118–142. Institute for Logic, Language and Computation, Universiteit van Amsterdam, Amsterdam, The Netherlands, 2010. URL: http://www.illc.uva.nl/lgc/seminar/?page_id=727. Cited on page **8**.

[72] P. Jirakunkanok, K. Sano, and S. Tojo. Analyzing belief re-revision by consideration of reliability change in legal case. In B. Mérialdo, M. L. Nguyen, D. Le, D. A. Duong, and S. Tojo, editors, *2015 Seventh International Conference on Knowledge and Systems Engineering, KSE 2015, Ho Chi Minh City, Vietnam, October 8-10, 2015*, pages 228–233. IEEE, 2015. ISBN 978-1-4673-8013-3. DOI: 10.1109/KSE.2015.64. URL: http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=7371541. Cited on page **35**.

[73] D. Katz and F. H. Allport. *Student Attitudes*. Craftsman, Syracuse, N.Y., 1931. Cited on page **2**.

[74] S. Konieczny and R. Pino Pérez. On the logic of merging. In A. G. Cohn, L. K. Schubert, and S. C. Shapiro, editors, *KR*, pages 488–498. Morgan Kaufmann, 1998. Cited on page **33**.

[75] S. Konieczny and R. Pino Pérez. Merging information under constraints: A logical framework. *Journal of Logic and Computation*, 12(5):773–808, 2002. DOI: 10.1093/logcom/12.5.773. Cited on pages **2 and 33**.

[76] S. Konieczny and R. Pino Pérez. Propositional belief base merging or how to merge beliefs/goals coming from several sources and some links with social choice theory. *European Journal of Operational Research*, 160(3):785–802, 2005. Cited on page **2**.

[77] S. Konieczny and R. Pino Pérez. Logic based merging. *Journal of Philosophical Logic*, 40(2):239–270, Mar. 2011. DOI: 10.1007/s10992-011-9175-5. Cited on pages **2 and 33**.

[78] D. Krech and R. S. Crutchfield. *Theory and Problems of Social Psychology*. McGraw-Hill, New York, 1948. Cited on page **2**.

[79] H. Lagerlund, S. Lindström, and R. Sliwinski, editors. *Modality Matters*. Number 53 in Uppsala Philosophical Studies. University of Uppsala, Upsala, 2006. Cited on pages **42 and 48**.

[80] I. H. LaValle and P. C. Fishburn. Lexicographic state-dependent subjective expected utility. *Journal of Risk and Uncertainty*, 4(3):251–269, July 1991. DOI: 10.1007/BF00114156. Cited on page **4**.

[81] I. H. LaValle and P. C. Fishburn. State-independent subjective expected lexicographic utility. *Journal of Risk and Uncertainty*, 5(3):217–240, July 1991. DOI: 10.1007/BF00057879. Cited on page **4**.

[82] K. Lehrer and C. Wagner. *Rational Consensus in Science and Society. A Philosophical and Mathematical Study*, volume 24 of *Philosophical Studies Series in Philosophy*. Dordrecht Reidel Publishing Company, Dordrecht, Holland, 1981. ISBN 978-90-277-1307-0. DOI: 10.1007/978-94-009-8520-9. Cited on pages **32 and 35**.

[83] H. Leibenstein. Bandwagon, snob, and veblen effects in the theory of consumers' demand. *Quarterly Journal of Economics*, 64(2):183–207, 1950. DOI: 10.2307/1882692. URL: http://www.jstor.org/stable/1882692. Cited on page 2.

[84] C.-J. Liau. Belief, information acquisition, and trust in multi-agent systems – a modal logic formulation. *Artificial Intelligence*, 149(1):31–60, 2003. DOI: 10.1016/S0004-3702(03)00063-8. Cited on page 8.

[85] V. Lifschitz. Some results on circumscription. Technical Report 1019, Stanford University, Dept of Computer Science, 1984. Cited on page 4.

[86] V. Lifschitz. Circumscription. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3: Nonmonotonic Reasoning and Uncertain Reasoning, pages 297–352. Oxford University Press, Oxford & New York, 1994. ISBN 0-19-853747-6. Cited on page 4.

[87] C. List and C. Puppe. Judgment aggregation. In P. Anand, P. Pattanaik, , and C. Puppe, editors, *The Handbook of Rational and Social Choice*. Oxford University Press, Oxford, May 2009. ISBN 978-0-199-29042-0. DOI: 10.1093/acprof:oso/9780199290420.003.0020. Cited on page 33.

[88] F. Liu. *Reasoning about Preference Dynamics*, volume 354 of *Synthese Library Series*. Springer, 2011. ISBN 978-94-007-1343-7. Cited on pages **2, 7, and 11**.

[89] F. Liu, J. Seligman, and P. Girard. Logical dynamics of belief change in the community. *Synthese*, 191(11):2403–2431, 2014. DOI: 10.1007/s11229-014-0432-3. Cited on page 33.

[90] E. Lorini and R. Demolombe. From binary trust to graded trust in information sources: A logical perspective. In Falcone et al. [45], pages 205–225. ISBN 978-3-540-92802-7. DOI: 10.1007/978-3-540-92803-4_11. Cited on page 8.

[91] E. Lorini, G. Jiang, and L. Perrussel. Trust-based belief change. In T. Schaub, G. Friedrich, and B. O'Sullivan, editors, *ECAI 2014 - 21st European Conference on Artificial Intelligence, 18-22 August 2014, Prague, Czech Republic - Including Prestigious Applications of Intelligent Systems (PAIS 2014)*, volume 263 of *Frontiers in Artificial Intelligence and Applications*, pages 549–554. IOS Press, 2014. DOI: 10.3233/978-1-61499-419-0-549. Cited on page 8.

[92] C. Lutz and U. Sattler. The complexity of reasoning with boolean modal logics. In F. Wolter, H. Wansing, M. de Rijke, and M. Zakharyaschev, editors, *Advances in Modal Logic 3, papers from the third conference on "Advances in Modal logic," held in Leipzig (Germany) in October 2000*, pages 329–348. World Scientific, 2000. ISBN 981-238-179-1. URL: www.informatik.uni-bremen.de/tdki/research/papers/2001/LutzSattler-AiML.ps.gz. Cited on page 9.

[93] D. van Mill. The possibility of rational outcomes from democratic discourse and procedures. *The Journal of Politics*, 58(3):734–752, Aug. 1996. DOI: 10.2307/2960440. Cited on page 2.

[94] D. Miller. Deliberative democracy and social choice. *Political Studies*, 40(s1):54–67, Aug. 1992. DOI: 10.1111/j.1467-9248.1992.tb01812.x. Cited on page 2.

[95] R. Nadeau, E. Cloutier, and J.-H. Guay. New evidence about the existence of a bandwagon effect in the opinion formation process. *International Political Science Review*, 14(2):203–213, 1993. DOI: 10.1177/019251219301400204. URL: http://www.jstor.org/stable/1882692. Cited on page 2.

[96] A. N. Prior. *Time and Modality*. Clarendon Press, Oxford, 1957. Cited on pages **9 and 13**.

[97] L. Ross, D. Greene, and P. House. The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13(3):279–301, may 1977. DOI: 10.1016/0022-1031(77)90049-X. Cited on page **2**.

[98] H. Rott. Shifting priorities: Simple representations for 27 iterated theory change operators. In Lagerlund et al. [79], pages 359–384. Also appeared as [99]. Cited on page **21**.

[99] H. Rott. Shifting priorities: Simple representations for twenty-seven iterated theory change operators. In D. Makinson, J. Malinowski, and H. Wansing, editors, *Towards Mathematical Philosophy*, volume 28 of *Trends in Logic*, pages 269–296. Springer Netherlands, 2009. ISBN 978-1-4020-9083-7. DOI: 10.1007/978-1-4020-9084-4_14. Cited on page **48**.

[100] M. D. Ryan. Defaults and revision in structured theories. In *Proceedings of the Sixth IEEE Symposium on Logic in Computer Science (LICS)*, pages 362–373, 1991. Cited on page **4**.

[101] M. D. Ryan. Belief revision and ordered theory presentations. In A. Fuhrmann and H. Rott, editors, *Logic, Action and Information*. De Gruyter Publishers, 1994. Also in Proceedings of the Eighth Amsterdam Colloquium on Logic, University of Amsterdam, 1991. Cited on page **4**.

[102] P.-Y. Schobbens. Exceptions for algebraic specifications: on the meaning of 'but'. *Science of Computer Programming*, 20(1-2):73–111, May 1993. DOI: 10.1016/0167-6423(93)90023-I. Cited on page **4**.

[103] Y. Wang and Q. Cao. On axiomatizations of public announcement logic. *Synthese*, 190(1):103–134, 2013. DOI: 10.1007/s11229-012-0233-5. Cited on page **24**.

[104] I. Welch. Sequential sales, learning and cascades. *Journal of Finance*, 47(2):695–732, 1992. ISSN 1540-6261. DOI: 10.1111/j.1540-6261.1992.tb04406.x. URL: `http://www.jstor.org/stable/2329120`. Cited on page **2**.

[105] G.-H. von Wright. *The logic of preference*. Edinburgh University Press, Edinburgh, 1963. Cited on page **7**.

[106] L. Zhen and J. Seligman. A logical model of the dynamics of peer pressure. *Electronic Notes in Theoretical Computer Science*, 278:275–288, 2011. DOI: 10.1016/j.entcs.2011.10.021. Cited on page **33**.