

A note on reliability-based preference dynamics

Sujata Ghosh¹ and Fernando R. Velázquez-Quesada²

¹ Indian Statistical Institute, Chennai, India. sujata@isichennai.res.in

² Grupo de Lógica, Lenguaje e Información, Universidad de Sevilla, Spain.
FRVelazquezQuesada@us.es

Abstract. This paper continues a line of work that studies individual preference upgrades in order to model situations akin to a process of public deliberation in collective decision making. It proposes a general upgrade policy, presenting its semantic definition and a corresponding modality for describing its effects as well as a complete axiom system.

1 Introduction

Deliberation and aggregation are essential and complementary components of any democratic decision making process. While the well-studied process of aggregation focuses on accumulating individual preferences without discussing their origin [1], deliberation can be seen as a conversation through which individuals justify their preferences, a process that might lead to changes in their opinions as they are influenced by one another. Even if deliberation does not lead to unanimity, the discussion can lead to some ‘preference uniformity’ (see how deliberation can help in bypassing social choice theory’s impossibility results in [2]), which might facilitate their eventual aggregation. In addition, the combination of both processes provides a more realistic model for decision making scenarios.

In [3], the authors presented a framework where agents have both preferences over a set of objects and reliability over the agents themselves. The main focus was to study how the public announcement of the individual preferences affects the preferences themselves. The paper proposed several preference upgrade policies based on the agents’ reliability orderings, and then introduced a general lexicographic upgrade operation subsuming all of them. Decision procedures were provided to decide whether, under such upgrade policies, the iterative and public announcement of individual preferences can eventually lead to preference unanimity/stability.

But not every ‘reasonable’ policy for upgrading individual preferences falls under the scope of the general lexicographic upgrade (see page 4 for a discussion) - this paper presents a more general upgrade policy, viz. the general layered upgrade. As we see in the example discussed later (cf. Example 1), the general definition provided in this paper captures intuitive upgrades which could not be formalised by policies discussed in [3]. Moreover, this short and technical note constitutes a necessary step towards formalizing reasonable deliberation processes which would facilitate their combination with aggregation processes in decision making. We leave this combination/reconciliation part for future work.

2 Recalling the framework

This section briefly recalls (and, in some cases, extends) the definitions of the *PR* framework; further details can be found in [3]. Throughout this paper, let Ag be a *finite non-empty* set of agents, with $|Ag| = n$.

Definition 1 (PR Frame). A preference and reliability (*PR*) frame F is a tuple $\langle W, \{\leq_i, \preceq_i\}_{i \in Ag} \rangle$ where (I) W is a finite non-empty set of worlds, (II) $\leq_i \subseteq (W \times W)$ is a total preorder (a total, reflexive and transitive relation), agent i 's preference relation over worlds in W ($u \leq_i v$ is read as “world v is at least as preferable as world u for agent i ”); (III) $\preceq_i \subseteq (Ag \times Ag)$ is a total order (a total, reflexive, transitive and antisymmetric relation), agent i 's reliability relation over agents in Ag ($j \preceq_i j'$ is read as “agent j' is at least as reliable as agent j for agent i ”).

The motivations for the restrictions on the preference and the reliability relations are discussed in [3]. For now, here are further useful definitions.

Definition 2. Let $F = \langle W, \{\leq_i, \preceq_i\}_{i \in Ag} \rangle$ be a frame.

- $\text{mr}(i) = j$ (j is agent i 's most reliable agent) iff_{def} $j' \preceq_i j$ for every $j' \in Ag$;
- $\text{Max}_{\leq_i}(U)$, the set containing agent i 's most preferred worlds among those in $U \subseteq W$, is formally defined as $\{v \in U \mid u \leq_i v \text{ for every } u \in U\}$.

2.1 A formal language

Throughout this paper, let At be a countable set of atomic propositions.

Definition 3 (Language). Formulas φ, ψ and relational expressions π, σ of the language \mathcal{L}^{PR} are given by

$$\begin{aligned} \varphi, \psi &::= \top \mid p \mid j \sqsubseteq_i j' \mid \neg\varphi \mid \varphi \vee \psi \mid \langle \pi \rangle \varphi \\ \pi, \sigma &::= 1 \mid \leq_i \mid \geq_i \mid ?(\varphi, \psi) \mid \neg\pi \mid \pi \cup \sigma \mid \pi \cap \sigma \end{aligned}$$

with $p \in At$ and $i, j, j' \in Ag$. Standard abbreviations as the converse operator $^{-1}$ over relational expressions³ will facilitate the writing of formulas.⁴

The set of formulas of \mathcal{L}^{PR} contains atomic propositions (p) and formulas describing the agents' reliability relations ($j \sqsubseteq_i j'$), 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), the preference relations (\leq_i), their respective converse (\geq_i ; [4,5]) and an additional construction of the form $?(\varphi, \psi)$ with φ and ψ formulas of the language, and it is closed under Boolean operations over relations (the so called *boolean modal logic*; [6]).

The following two definitions establish what a model is and how formulas of \mathcal{L}^{PR} are interpreted over such structures.

³ Such operator is given by $1^{-1} := 1$, $(\leq_i)^{-1} := \geq_i$, $(\geq_i)^{-1} := \leq_i$, $(?(\varphi, \psi))^{-1} := ?(\psi, \varphi)$, $(\neg\pi)^{-1} := \neg(\pi^{-1})$, $(\pi \cup \sigma)^{-1} := \pi^{-1} \cup \sigma^{-1}$ and $(\pi \cap \sigma)^{-1} := \pi^{-1} \cap \sigma^{-1}$.

⁴ Additionally, $\langle <_i \rangle \varphi := \langle \leq_i \cap \neg \geq_i \rangle \varphi$ and $\langle >_i \rangle \varphi := \langle \neg \leq_i \cap \geq_i \rangle \varphi$.

Definition 4 (PR model). A PR model M is a tuple $\langle F, V \rangle$ where F is a PR frame and $V : At \rightarrow \wp(W)$ is a valuation function.

Definition 5 (Semantic interpretation). Let $M = \langle W, \{\leq_i, \preceq_i\}_{i \in Ag}, V \rangle$ be a PR model. The function $\llbracket \cdot \rrbracket^M$ from formulas in \mathcal{L}^{PR} to subsets of W and the function $\lceil \cdot \rceil^M$ from relational expressions in \mathcal{L}^{PR} to binary relations over W are defined simultaneously in the following way.

$$\begin{aligned} \llbracket \top \rrbracket^M &:= W & \llbracket p \rrbracket^M &:= V(p) & \llbracket j \sqsubseteq_i j' \rrbracket^M &:= \begin{cases} W & \text{if } j \preceq_i j' \\ \emptyset & \text{otherwise} \end{cases} \\ \llbracket \neg \varphi \rrbracket^M &:= W \setminus \llbracket \varphi \rrbracket^M & \llbracket \varphi \vee \psi \rrbracket^M &:= \llbracket \varphi \rrbracket^M \cup \llbracket \psi \rrbracket^M \\ \llbracket \langle \pi \rangle \varphi \rrbracket^M &:= \{w \in W \mid \text{there is } u \in \llbracket \varphi \rrbracket^M \text{ with } (w, u) \in \lceil \pi \rceil^M\} \end{aligned}$$

and

$$\begin{aligned} \lceil 1 \rceil^M &:= W \times W & \lceil \neg \pi \rceil^M &:= (W \times W) \setminus \lceil \pi \rceil^M \\ \lceil \leq_i \rceil^M &:= \leq_i & \lceil \pi \cup \sigma \rceil^M &:= \lceil \pi \rceil^M \cup \lceil \sigma \rceil^M \\ \lceil \geq_i \rceil^M &:= \{(v, u) \in (W \times W) \mid u \leq_i v\} & \lceil \pi \cap \sigma \rceil^M &:= \lceil \pi \rceil^M \cap \lceil \sigma \rceil^M \\ \lceil ?(\varphi, \psi) \rceil^M &:= \llbracket \varphi \rrbracket^M \times \llbracket \psi \rrbracket^M \end{aligned}$$

Note, in particular, how $\lceil ?(\varphi, \psi) \rceil^M$ is the set of those pairs $(u, v) \in (W \times W)$ such that u satisfies φ and v satisfies ψ .⁵ A formula φ is true at world $w \in W$ in model M when $w \in \llbracket \varphi \rrbracket^M$. A formula is valid when it is true at every world of every model, as usual.

The operator $?(\varphi, \psi)$, useful for providing the axiom system for the general upgrade operation to be introduced in Subsection 3.1, is the only construction in \mathcal{L}^{PR} that does not appear in [3]. Thus, an axiom system characterising formulas in \mathcal{L}^{PR} valid on PR models is given by the axioms and rules in Table 1 of [3] plus the formula $\langle ?(\psi_1, \psi_2) \rangle \varphi \leftrightarrow (\psi_1 \wedge \langle 1 \rangle (\psi_2 \wedge \varphi))$, which characterises the extra operator.

3 Individual preference upgrades

Intuitively, a public announcement of the agents' individual preferences might induce an agent i to adjust her own preferences according to what has been announced and the reliability ordering she assigns to the set of agents.⁶ For example, an agent might adopt the preferences of the agent on whom she relies the most, or might use such preference for 'breaking ties' among her equally-preferred zones.

In [3] the authors introduced the *general lexicographic upgrade* operation, which creates a preference ordering following a priority list of orderings.

⁵ The relation $\lceil ?(\varphi, \psi) \rceil^M$ is a natural generalisation of the relation $\lceil ?\varphi \rceil^M := \{(u, u) \in (W \times W) \mid u \in \llbracket \varphi \rrbracket^M\}$ for the traditional PDL test operation $? \varphi$ [7].

⁶ Note that this work does not focus on the formal representation of such announcement, but rather on the formal representation of its effects.

Definition 6 (General lexicographic upgrade). A lexicographic list \mathcal{R} over W is a finite non-empty list whose elements are indexes of preference orderings over W , with $|\mathcal{R}|$ the list's length and $\mathcal{R}[k]$ its k th element ($1 \leq k \leq |\mathcal{R}|$). Intuitively, \mathcal{R} is a priority list of preference orderings, with $\leq_{\mathcal{R}[1]}$ having the highest priority. Given \mathcal{R} , the preference ordering $\leq_{\mathcal{R}} \subseteq (W \times W)$ is defined as

$$u \leq_{\mathcal{R}} v \quad \text{iff}_{def} \quad \underbrace{\left(u \leq_{\mathcal{R}[|\mathcal{R}|]} v \wedge \bigwedge_{k=1}^{|\mathcal{R}|-1} u \simeq_{\mathcal{R}[k]} v \right)}_1 \vee \underbrace{\left(\bigvee_{k=1}^{|\mathcal{R}|-1} \left(u <_{\mathcal{R}[k]} v \wedge \bigwedge_{l=1}^{k-1} u \simeq_{\mathcal{R}[l]} v \right) \right)}_2$$

Thus, $u \leq_{\mathcal{R}} v$ holds if this agrees with the least prioritised ordering ($\leq_{\mathcal{R}[|\mathcal{R}|]}$) and for the rest of them u and v are equally preferred (part 1), or if there is an ordering $\leq_{\mathcal{R}[k]}$ with a strict preference for v over u and all orderings with higher priority see u and v as equally preferred (part 2).

This operation allows an agent i to upgrade her preferences by taking $\leq'_i := \leq_{\mathcal{R}}$, with \mathcal{R} a lexicographic list containing the ordered indexes of the agents whose preferences will be used. It subsumes not only the natural instance in which \mathcal{R} is given directly by the agent's reliability ordering, but also other possibilities as, e.g., one in which the agent adopts 'as is' the preferences of the agent on whom she relies the most. A sound and complete axiom system for a modality representing the operation can be found in [3].

Even though the general lexicographic upgrade covers many natural upgrades, there are also 'reasonable' policies that fall outside its scope. Sometimes we are not interested in considering the complete order among the choices of the most reliable agent, but only her most preferred choices. For example, consider a girl planning to take a boy out for a movie of his choice, and let w_i denote the world where 'movie i is the most preferred' ($i = 1, 2, 3, 4$). Rather than considering the complete preference ordering it makes more sense to consider the most preferred movies of the boy and among that what she would like to watch most as well. In any case they can take note of their choices among all the options, as they may not know which movie ticket they would get. We will get back to this example in a moment. For the following definition, recall that $\text{Max}_{\leq_i}(W)$ denotes agent i 's most preferred worlds among those in W .

Definition 7 (Conservative upgrade). Agent j put her most reliable agent's most preferred worlds above the rest, using her old ordering to break ties in both zones.⁷ More precisely, with $U := \text{Max}_{\leq_{\text{mr}(j)}}(W)$,

$$u \leq'_j v \quad \text{iff}_{def} \quad (\{u, v\} \cap U = \{u, v\} \wedge u \leq_j v) \vee (\{u, v\} \cap U = \{v\}) \vee (\{u, v\} \cap U = \emptyset \wedge u \leq_j v)$$

The conservative upgrade is not an instance of the general lexicographic upgrade, as there are cases in which the output of the former cannot be reproduced by any instance of the latter.

⁷ This upgrade is called *lexicographic* in [8] and [9].

Example 1. Suppose agent a is agent b 's most reliable agent and their individual preferences are as below (reflexive and transitive arrows omitted).



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 producing $w_3 <'_b w_4$ and $w_1 <'_b w_2$. The final result is then

$$b: w_1 \longrightarrow w_2 \longrightarrow w_3 \longrightarrow w_4$$

Observe how no lexicographic list can produce this outcome. First, no singleton list does the job, as \leq'_b is different from both \leq_a and \leq_b . The list $\langle\langle a ; b \rangle\rangle$ (with the leftmost ordering having the highest priority) also fails, as it would give \leq_a the highest priority, thus producing an ordering with w_2 strictly below w_1 , different from what \leq'_b states. Finally, $\langle\langle b ; a \rangle\rangle$ fails too, as it will give priority to \leq_b , thus putting w_4 strictly below w_1 , again different from what \leq'_b establishes.

Now agents a and b can be considered as the boy and girl respectively in the earlier example, with their preference orders about movies given as above. After a conservative preference upgrade, the first choice movie for the girl is movie 4.

3.1 The general layered upgrade

The conservative upgrade does not create a preference ordering following a priority list of orderings. Instead, it puts a set of elements of the domain at the topmost layer of the ordering (in Definition 7, the set $\text{Max}_{\leq_{\text{mr}(j)}}(W)$), then using a ‘default’ ordering (in Definition 7, \leq_j) to sort both this layer and those worlds that do not appear in it. This observation leads to the following definition.

Definition 8 (General layered upgrade). A layered list \mathcal{S} over W is a finite (possibly empty) list of pairwise disjoint subsets of W together with a default preference ordering over W . The list's length is denoted by $|\mathcal{S}|$, its k th element is denoted by $\mathcal{S}[k]$ (with $1 \leq k \leq |\mathcal{S}|$), and $\leq_{\text{def}}^{\mathcal{S}}$ is its default preference ordering. Intuitively, \mathcal{S} defines layers of elements of W in the new preference ordering $\leq_{\mathcal{S}}$, with $\mathcal{S}[1]$ the set of worlds that will be in the topmost layer and $\leq_{\text{def}}^{\mathcal{S}}$ the preference ordering that will be applied to each individual set and to those worlds not in $\bigcup_{k=1}^{|\mathcal{S}|} \mathcal{S}[k]$. Formally, given \mathcal{S} , the ordering $\leq_{\mathcal{S}} \subseteq (W \times W)$ is defined as

$$\begin{aligned}
 u \leq_{\mathcal{S}} v \quad \text{iff}_{\text{def}} \quad & \underbrace{\left(u \leq_{\text{def}}^{\mathcal{S}} v \wedge \left(\{u, v\} \cap \bigcup_{k=1}^{|\mathcal{S}|} \mathcal{S}[k] = \emptyset \vee \bigvee_{k=1}^{|\mathcal{S}|} \{u, v\} \subseteq \mathcal{S}[k] \right) \right)}_1 \\
 \vee \quad & \underbrace{\bigvee_{k=1}^{|\mathcal{S}|} \left(v \in \mathcal{S}[k] \wedge u \notin \bigcup_{l=1}^k \mathcal{S}[l] \right)}_2
 \end{aligned}$$

Thus, $u \leq_S v$ holds if this agrees with the default ordering \leq_{def}^S and either neither u nor v are in any of the specified sets in \mathcal{S} or else both are in the same set (part 1), or if there is a set $\mathcal{S}[k]$ in which v appears and u appears neither in the same set (a case already covered in part 1) nor in one with higher priority (part 2).

Here are two useful observations. First, if $|\mathcal{S}| = 0$, then while both the whole part 2 and the right-hand side of the rightmost disjunct in part 1 collapse to \perp , the left-hand side of the rightmost disjunct in part 1 collapses to \top . Thus,

$$u \leq_S v \quad \text{iff} \quad u \leq_{\text{def}}^S v$$

On the other hand, if \mathcal{S} 's sets form a partition of W (i.e., the sets are not only mutually exclusive but also collectively exhaustive), then $\bigcup_{k=1}^{|\mathcal{S}|} \mathcal{S}[k] = W$ so the left-hand side of the rightmost disjunct in part 1 collapses to \perp . Then,

$$u \leq_S v \quad \text{iff} \quad \underbrace{\left(u \leq_{\text{def}}^S v \wedge \bigvee_{k=1}^{|\mathcal{S}|} \{u, v\} \subseteq \mathcal{S}[k] \right)}_1 \vee \underbrace{\bigvee_{k=1}^{|\mathcal{S}|} \left(v \in \mathcal{S}[k] \wedge u \notin \bigcup_{l=1}^k \mathcal{S}[l] \right)}_2$$

In fact, since \leq_{def}^S is used to break ties not only within each $\mathcal{S}[k]$ but also among those worlds not appearing in any such set, the provided definition of a layered list actually just ‘abbreviates’ (but still it is equivalent to) a list that requires a full partition of W by not writing explicitly the set with the least priority.

Third, a layered list \mathcal{S} has a semantic nature, as it is given in terms of subsets of the domain and binary relations over it. Of course, when it is intended to be applied to a given model, it can also be defined syntactically.

Definition 9. A layered list \mathcal{S} is defined syntactically within \mathcal{L}^{PR} whenever each $\mathcal{S}[k]$ is given as a formula χ_k in \mathcal{L}^{PR} and its default ordering \leq_{def}^S is given as a relational expression π_{def}^S in \mathcal{L}^{PR} . In such cases, Definition 8 is adjusted by writing $[\chi_k]^M$ instead of $\mathcal{S}[k]$ and $\ulcorner \pi_{\text{def}}^S \urcorner^M$ instead of \leq_{def}^S , for M the model in which such layered list is applied. In such cases, \leq_S will be written as $\leq_{S(M)}$.

The next proposition makes possible the definition that follows it.

Proposition 1. Let \mathcal{S} be a layered list over W . If \leq_{def}^S is reflexive (transitive, total, respectively), then so is \leq_S . For syntactically defined layered lists \mathcal{S} , if $\ulcorner \pi_{\text{def}}^S \urcorner^M$ is reflexive (transitive, total, respectively), then so is $\leq_{S(M)}$.

Definition 10. Let $M = \langle W, \{\leq_i, \preceq_i\}_{i \in \text{Ag}}, V \rangle$ be a PR model.

- Let \mathcal{S} be a layered list whose default ordering is reflexive, transitive and total;⁸ let $j \in \text{Ag}$ be an agent. The PR model $\text{gy}_{\mathcal{S}}^j(M) = \langle W, \{\leq'_i, \preceq'_i\}_{i \in \text{Ag}}, V \rangle$ is such that, for every agent $i \in \text{Ag}$, $\leq'_i := \leq_{S(M)}$ if $i = j$, and $\leq'_i := \leq_i$ otherwise.
- Let \mathcal{S} be a list of $|\text{Ag}|$ layered lists whose default ordering are reflexive, transitive and total, with \mathcal{S}_i its i th element.⁹ The PR model $\text{gy}_{\mathcal{S}}(M) = \langle W, \{\leq'_i, \preceq'_i\}_{i \in \text{Ag}}, V \rangle$ is such that, for every agent $i \in \text{Ag}$, $\leq'_i := \leq_{\mathcal{S}_i(M)}$

⁸ If \mathcal{S} is defined syntactically, then $\ulcorner \pi_{\text{def}}^S \urcorner^M$ should be reflexive, transitive and total.

⁹ If \mathcal{S}_i is defined syntactically, then $\ulcorner \pi_{\text{def}}^{\mathcal{S}_i} \urcorner^M$ should be reflexive, transitive and total.

On the generality of the general layered upgrade When layered lists are defined semantically, the general layered upgrade can build *any* conceivable *total, reflexive and transitive* preference ordering by simply using a layered list that spells out explicitly the desired output, using then the full Cartesian product as the default ordering.

When layered lists are restricted to syntactically definable ones, the power of the layered upgrade depends on the expressivity of the used language. Nevertheless, \mathcal{L}^{PR} is expressive enough to define layered lists that replicate the behaviour of not only the general lexicographic (Definition 6) but also the conservative upgrade (Definition 7). This shows how the general layered upgrade is indeed a generalisation of the general lexicographic upgrade.

Proposition 2. *The general lexicographic upgrade is an instance of the general layered upgrade with \mathcal{S} defined syntactically within \mathcal{L}^{PR} .*

Proof (Sketch). Let $M = \langle W, \{\leq_i, \preceq_i\}_{i \in Ag}, V \rangle$ be a PR model. Take any lexicographic list \mathcal{R} , and let L_1, \dots, L_m be the layers it generates (with L_1 being the topmost) when applied over M . If the relational expression 1 is used for defining the default ordering, then in order to prove the proposition it is enough to provide m formulas χ_k such that $L_k = \llbracket \chi_k \rrbracket^M$. In order to do this, first observe how, if $U = \llbracket \chi_U \rrbracket^M$, then $\text{Max}_{\leq_i}(U) = \llbracket \chi_U \wedge [\leq_i] \neg \chi_U \rrbracket^M$. Now, note how

$$L_1 = \text{Max}_{\leq_{\mathcal{R}[|\mathcal{R}|]}} \left(\text{Max}_{\leq_{\mathcal{R}[|\mathcal{R}|-1]}} \left(\dots \text{Max}_{\leq_{\mathcal{R}[1]}}(W) \dots \right) \right)$$

This and the previous observation suggest the following recursive definition:

$$\begin{aligned} \mu_1(\tau) &:= \tau \wedge [\leq_{\mathcal{R}[1]}] \neg \tau \\ \mu_2(\tau) &:= \mu_1(\tau) \wedge [\leq_{\mathcal{R}[2]}] \neg \mu_1(\tau) \\ &\vdots \\ \mu_{|\mathcal{R}|}(\tau) &:= \mu_{|\mathcal{R}|-1}(\tau) \wedge [\leq_{|\mathcal{R}|}] \neg \mu_{|\mathcal{R}|-1}(\tau) \end{aligned}$$

in which τ is a parameter. Then, given $W = \llbracket \top \rrbracket^M$, it is not hard to see that

$$\chi_1 := \mu_{|\mathcal{R}|}(\top) \quad \text{is such that} \quad L_1 = \llbracket \chi_1 \rrbracket^M$$

But then, since

$$L_2 = \text{Max}_{\leq_{\mathcal{R}[|\mathcal{R}|]}} \left(\text{Max}_{\leq_{\mathcal{R}[|\mathcal{R}|-1]}} \left(\dots \text{Max}_{\leq_{\mathcal{R}[1]}}(W \setminus L_1) \dots \right) \right)$$

it follows that

$$\chi_2 := \mu_{|\mathcal{R}|}(\top \wedge \neg \chi_1) \quad \text{is such that} \quad L_2 = \llbracket \chi_2 \rrbracket^M$$

This process can be repeated. In its m th iteration, by observing

$$L_m = \text{Max}_{\leq_{\mathcal{R}[|\mathcal{R}|]}} \left(\text{Max}_{\leq_{\mathcal{R}[|\mathcal{R}|-1]}} \left(\dots \text{Max}_{\leq_{\mathcal{R}[1]}}(W \setminus \bigcup_{k=1}^{m-1} L_k) \dots \right) \right)$$

it follows that

$$\chi_m := \mu_{|\mathcal{R}|}(\top \wedge \bigwedge_{k=1}^{m-1} \neg \chi_k) \quad \text{is such that} \quad L_m = \llbracket \chi_m \rrbracket^M$$

The process stops here, as $W = L_1 \cup \dots \cup L_m$, and thus further iterations will produce formulas χ such that $\llbracket \chi \rrbracket^M = \emptyset$.

Proposition 3. *The conservative upgrade is an instance of the general layered upgrade with \mathcal{S} defined syntactically within \mathcal{L}^{PR} .*

Proof (Sketch). It is enough to provide the explicit definition of a syntactically defined layered list that does the job. Take any PR model M with domain W and observe how $\llbracket \bigwedge_{i' \in Ag} (i' \sqsubseteq_j i) \rrbracket^M = W$ iff $i = \text{mr}(j)$. Then,

$$\chi := \bigwedge_{i' \in Ag} (i' \sqsubseteq_j i) \rightarrow [\langle i \rangle \perp \quad \text{implies} \quad \text{Max}_{\leq_{\text{mr}(j)}}(W) = \llbracket \chi \rrbracket^M$$

Thus, a ‘singleton list’ with χ as its unique set and \leq_j as its default relational expression induces the ordering generated by the conservative upgrade.

A more illuminating way to prove how the general layered upgrade indeed extends the general lexicographic one is by noticing that, while the general lexicographic upgrade cannot revert strict preferences when these are unanimous, the general layered can. More precisely, on the one hand,

Proposition 4. *For every PR model, if all agents put a given world strictly above another, then so does the ordering $\leq_{\mathcal{R}}$ induced by any lexicographic \mathcal{R} .*

Nevertheless, on the other hand,

Fact 1. *There are PR models in which all agents agree in the relative strict order between two worlds, and yet a general layered upgrade can reverse it.*

Proof. Take a frame with a single agent having a strict preference of w_1 over w_2 . This order is switched by using a general layered upgrade with a singleton list given by $[\rangle] \perp$ and with 1 being the relational expression for its default ordering.

The generality offered by the general layered upgrade might be welcomed from some perspectives, but it might not be completely desirable from others. For example, when the layered list is given syntactically by \mathcal{L}^{PR} , it allows the definition of ‘unreasonable’ preference upgrade policies. As an illustration, a singleton layered list for agent j with its set defined by $[\rangle_{\text{mr}(j)}] \perp$ (with $\text{mr}(j)$ characterised as in the proof of Proposition 3) will move to the top of j ’s ordering those worlds that are, for her *most reliable* agent, the *least preferred* ones.

From this perspective, the general lexicographic upgrade has some advantages over the layered one: not only combines the current orderings in a ‘natural’ way, but also has some pleasant properties, as it respects unanimity not only over strict preferences, as shown before, but also over equal-preferability.

Proposition 5. *For every PR model, if all agents agree in that two worlds are equally preferred, then so does the ordering $\leq_{\mathcal{R}}$ induced by any lexicographic \mathcal{R} .*

A more detailed study of the expressivity of the general layered upgrade when the used list is syntactically defined within \mathcal{L}^{PR} is left for future work.

Some observations and a comparison In the literature one can find several operations describing changes in orderings among objects. In particular, there are several *dynamic epistemic logic* [10,11] proposals for orderings interpreted not only as preferences (so the operations represent preference change: e.g., [12,13]), but also as plausibility (so the operations are understood as forms of belief revision: e.g., [8,9,14,15]). It is worthwhile to discuss, albeit briefly, some key characteristics of the general layered upgrade and how it relates to existing frameworks.

A straightforward observation is that the general layered upgrade (*GLay*) only affects the ordering, keeping the domain intact (thus differing from, e.g., [14,15]). More interesting is the fact that, although it generalises the general lexicographic upgrade (*GLex*) of [3], *GLay* still has a lexicographic spirit: the sets in \mathcal{S} actually define an ordering, and thus $\leq_{\mathcal{S}}$ is the result of a lexicographic upgrade with the order generated by the sets having the highest priority, and the default ordering being used only to ‘break ties’.

A closer comparison between *GLay* and the plausibility action models (*PAM*) of [15] is also useful. They share the same spirit, as a *PAM* is a relational structure in which each ordered ‘world’ is associated to a formula, thus defining in this way an ordering among sets of worlds, just as the sets of a layered list in *GLay*. Moreover, a *PAM* acts over a plausibility model following the ‘action priority’ rule: the ordering in the resulting model is a combination of the one in initial model with that of the *PAM* in which the latter has the priority, exactly as *GLay* does when it prioritises the sets over the default ordering. In fact, the crucial difference between these frameworks might be simply the expressivity of the language used for both the initial ordering and the default one. With respect to the initial ordering, the language used in the *PAM* framework is equivalent to the fragment of \mathcal{L}^{PR} in which the relational expressions are only \leq_i and $\sim_i := \leq_i \cup \geq_i$; maybe more important, with respect to the default ordering, by construction *PAM* uses only the agent’s preference relation, but *GLay* allows a full relational expression. Whether this difference in expressivity allows *GLay* to create orderings that cannot be defined by using *PAM* remains to be studied.

Note that the layers of a ‘layered list’ can be interpreted as levels of beliefs in such plausibility models, and also in the *KD45-O* models of [16]. The basic difference here is the additional ‘default ordering’ which makes sense while describing preferences as it can be thought of as some preference ordering exogenously instilled in agents, which comes to the fore when absolutely necessary.

The formal language

Definition 11. *The language $\mathcal{L}_{\{\text{gy}\}}^{PR}$ extends \mathcal{L}^{PR} with a modality $\langle \text{gy}_{\mathcal{S}}^i \rangle$ for every agent $i \in \text{Ag}$ and every layered list \mathcal{S} whose default ordering is reflexive, transitive and total. Given a PR model M , define*

$$\llbracket \langle \text{gy}_{\mathcal{S}}^i \rangle \varphi \rrbracket^M := \llbracket \varphi \rrbracket^{\text{gy}_{\mathcal{S}}^i(M)}$$

with $\text{gy}_{\mathcal{S}}^i(M)$ as in Definition 10. Note how, by defining $[\text{gy}_{\mathcal{S}}^i] \varphi := \neg \langle \text{gy}_{\mathcal{S}}^i \rangle \neg \varphi$, then $\llbracket [\text{gy}_{\mathcal{S}}^i] \varphi \rrbracket^M := \llbracket \varphi \rrbracket^{\text{gy}_{\mathcal{S}}^i(M)}$ so $\langle \text{gy}_{\mathcal{S}}^i \rangle \varphi \leftrightarrow [\text{gy}_{\mathcal{S}}^i] \varphi$ is valid.

The modality $\langle \text{gy}_{\mathcal{S}}^i \rangle$ allows to describe the effects of upgrading agent i 's preferences via the general layered upgrade with \mathcal{S} , keeping the preferences of the remaining agents as before. This definition can be extended to simultaneous upgrades by asking for a list \mathcal{S} of layered lists and using a modality $\langle \text{gy}_{\mathcal{S}} \rangle$ whose semantic interpretation uses the operation $\text{gy}_{\mathcal{S}}(\cdot)$ of Definition 10.

For an axiom system, this paper provides valid formulas and validity-preserving rules indicating how to rewrite a formula using $\langle \text{gy}_{\mathcal{S}}^i \rangle$ as a provably equivalent one in \mathcal{L}^{PR} . Then, while soundness follows from the validity and validity preserving properties of the rewriting tools, completeness follows from the completeness of the basic ‘static’ system (end of Subsection 2.1).¹⁰

Besides indicating how to translate atomic propositions, reliability formulas and their Boolean combinations, the rewriting formulas should indicate how to translate formulas involving modal operators of the form $\langle \pi \rangle$, where π can be any relational expression. Hence, given any relational expression in the model $\text{gy}_{\mathcal{S}}^i(M)$, a ‘matching’ relational expression in the original model M should be provided. The layered relational transformer defined below, similar in spirit to the program transformers of [18] for providing rewriting axioms for regular *PDL*-expressions [7] (in their case, after the action-model updates of [19]), will capture this. However, in order to express within \mathcal{L}^{PR} the effect of a general layered upgrade, the used layered list \mathcal{S} must be syntactically defined in \mathcal{L}^{PR} : indeed, if either some $\mathcal{S}[k]$ or else the default ordering $\leq_{\text{def}}^{\mathcal{S}}$ is not \mathcal{L}^{PR} -definable, then the language cannot tell whether a world is in $\mathcal{S}[k]$ or whether a pair satisfies $\leq_{\text{def}}^{\mathcal{S}}$, and thus it cannot describe the upgrade’s effects.

Definition 12 (Layered relational transformer). *Let M be a PR model with domain W ; let i be an agent. Let \mathcal{S} be a syntactically defined layered list over W for which each set $\mathcal{S}[k]$ is characterised by a formula $\chi_{\mathcal{S}[k]}$ in \mathcal{L}^{PR} (i.e., $\mathcal{S}[k] = \llbracket \chi_{\mathcal{S}[k]} \rrbracket^M$) and whose default ordering $\leq_{\text{def}}^{\mathcal{S}}$ is characterised by a relational expression $\pi_{\text{def}}^{\mathcal{S}}$ in \mathcal{L}^{PR} (i.e., $\leq_{\text{def}}^{\mathcal{S}} = \llbracket \pi_{\text{def}}^{\mathcal{S}} \rrbracket^M$). Define $N_{\mathcal{S}}$ as the formula satisfied by those worlds that do not appear in a set in \mathcal{S} , and $N_{\mathcal{S}}^k$ as the formula satisfied by those worlds that do not appear in the sets $\mathcal{S}[1], \dots, \mathcal{S}[k]$ for some index k .¹¹*

$$N_{\mathcal{S}} := \neg \bigvee_{k=1}^{|\mathcal{S}|} \chi_{\mathcal{S}[k]} \quad N_{\mathcal{S}}^k := \neg \bigvee_{l=1}^k \chi_{\mathcal{S}[l]}$$

A layered relational transformer $Ty_{\mathcal{S}}^i$ is a function from relational expressions to relational expressions defined in the following way.

$$Ty_{\mathcal{S}}^i(\leq_i) := \underbrace{\left(\pi_{\text{def}}^{\mathcal{S}} \cap \left(?(N_{\mathcal{S}}, N_{\mathcal{S}}) \cup \bigcup_{k=1}^{|\mathcal{S}|} ?(\chi_{\mathcal{S}[k]}, \chi_{\mathcal{S}[k]}) \right) \right)}_1 \cup \underbrace{\bigcup_{k=1}^{|\mathcal{S}|} ?(N_{\mathcal{S}}^k, \chi_{\mathcal{S}[k]})}_2$$

$$Ty_{\mathcal{S}}^i(\geq_i) := \underbrace{\left((\pi_{\text{def}}^{\mathcal{S}})^{-1} \cap \left(?(N_{\mathcal{S}}, N_{\mathcal{S}}) \cup \bigcup_{k=1}^{|\mathcal{S}|} ?(\chi_{\mathcal{S}[k]}, \chi_{\mathcal{S}[k]}) \right) \right)}_1 \cup \underbrace{\bigcup_{k=1}^{|\mathcal{S}|} ?(\chi_{\mathcal{S}[k]}, N_{\mathcal{S}}^k)}_2$$

¹⁰ See Chapter 7 of [10] (cf. [17]) for an extensive explanation of this technique.

¹¹ The case with $|\mathcal{S}| = 0$ can be understood as a case in which each $\mathcal{S}[k]$ is the empty set, and thus $\chi_{\mathcal{S}[k]} = \perp$. In such case, $N_{\mathcal{S}}$ becomes the always true \top .

$\vdash \langle \text{gy}_S^i \rangle \top$	$\vdash \langle \text{gy}_S^i \rangle (\varphi \vee \psi) \leftrightarrow (\langle \text{gy}_S^i \rangle \varphi \vee \langle \text{gy}_S^i \rangle \psi)$
$\vdash \langle \text{gy}_S^i \rangle p \leftrightarrow p$	$\vdash \langle \text{gy}_S^i \rangle (\varphi \rightarrow \psi) \leftrightarrow (\langle \text{gy}_S^i \rangle \varphi \rightarrow \langle \text{gy}_S^i \rangle \psi)$
$\vdash \langle \text{gy}_S^i \rangle j' \sqsubseteq_j j'' \leftrightarrow j' \sqsubseteq_j j''$	$\vdash \langle \text{gy}_S^i \rangle \langle \pi \rangle \varphi \leftrightarrow \langle \text{Ty}_S^i(\pi) \rangle \langle \text{gy}_S^i \rangle \varphi$
$\vdash \langle \text{gy}_S^i \rangle \neg \varphi \leftrightarrow \neg \langle \text{gy}_S^i \rangle \varphi$	From $\vdash \varphi$ infer $\vdash [\text{gy}_S^i] \varphi$

Table 1. Axioms for $\mathcal{L}_{\{\text{gy}\}}^{PR}$ w.r.t. PR models.

$$\begin{aligned}
\text{Ty}_S^i(1) &:= 1 & \text{Ty}_S^i(-\pi) &:= -\text{Ty}_S^i(\pi) \\
\text{Ty}_S^i(\leq_j) &:= \leq_j \quad \text{for } i \neq j & \text{Ty}_S^i(\pi \cup \sigma) &:= \text{Ty}_S^i(\pi) \cup \text{Ty}_S^i(\sigma) \\
\text{Ty}_S^i(\geq_j) &:= \geq_j \quad \text{for } i \neq j & \text{Ty}_S^i(\pi \cap \sigma) &:= \text{Ty}_S^i(\pi) \cap \text{Ty}_S^i(\sigma) \\
\text{Ty}_S^i(?(\psi_1, \psi_2)) &:= ?(\langle \text{gy}_S^i \rangle \psi_1, \langle \text{gy}_S^i \rangle \psi_2)
\end{aligned}$$

Intuitively, a layered relational transformer Ty_S^i takes a relational expression representing a relation in the model $\text{gy}_S^i(M)$ and returns a matching relational expression representing a relation in the original model M . The cases for the basic relational expressions, \leq_i and \geq_i , are the important ones. The first uses Definition 8 to establish that \leq_i in $\text{gy}_S^i(M)$ corresponds to $\leq_{S(M)}$ in M ; the second uses the same definition to indicate that \geq_i in $\text{gy}_S^i(M)$ is the converse of $\leq_{S(M)}$ in M . The remaining cases take care of the constant 1, the basic relational expressions for agents other than i and of the relational test as well as the complement, union and intersection of relations. With Ty_S^i defined, it is possible now to provide the promised axiom system.

Theorem 2. *The axioms and rules on Table 1 together with those of the basic ‘static’ system (end of Subsection 2.1) provide a sound and complete axiom system (with i any agent) for $\mathcal{L}_{\{\text{gy}\}}^{PR}$ with respect to PR models.*

Proof (Sketch). The rule and the axioms for atomic propositions, reliability, negation and disjunction are standard for an operation without precondition that does not affect atomic propositions (and, in this case, neither reliability). The axiom for relational expressions, the key one, makes crucial use of the layered relational transformer, stating that there is a φ -world π -reachable from the evaluation point at $\text{gy}_S^i(M)$ if and only if there is a $\langle \text{gy}_S^i \rangle \varphi$ -world $\text{Ty}_S^i(\pi)$ -reachable from the evaluation point at M . As an example, if π is \leq_i , then the axiom is

$$\langle \text{gy}_S^i \rangle \langle \leq_i \rangle \varphi \leftrightarrow \underbrace{\left(\pi_{\text{def}}^S \cap \left(?(N_S, N_S) \cup \bigcup_{k=1}^{|\mathcal{S}|} ?(\chi_{S[k]}, \chi_{S[k]}) \right) \right)}_1 \cup \underbrace{\bigcup_{k=1}^{|\mathcal{S}|} ?(N_S^k, \chi_{S[k]})}_{2} \langle \text{gy}_S^i \rangle \varphi$$

whose right-hand side, by using the axioms for \cup and $?$ together with some commutation and distribution, is equivalent to

$$\begin{aligned}
& \underbrace{\langle \pi_{\text{def}}^{\mathcal{S}} \cap ?(N_{\mathcal{S}}, N_{\mathcal{S}}) \rangle \langle \text{gy}_{\mathcal{S}}^i \rangle \varphi \vee \bigvee_{k=1}^{|\mathcal{S}|} \langle \pi_{\text{def}}^{\mathcal{S}} \cap ?(\chi_{\mathcal{S}[k]}, \chi_{\mathcal{S}[k]}) \rangle \langle \text{gy}_{\mathcal{S}}^i \rangle \varphi}_{1} \\
& \vee \underbrace{\bigvee_{k=1}^{|\mathcal{S}|} \left(N_{\mathcal{S}}^k \wedge \langle 1 \rangle (\chi_{\mathcal{S}[k]} \wedge \langle \text{gy}_{\mathcal{S}}^i \rangle \varphi) \right)}_{2}
\end{aligned}$$

Thus, the axiom states that after a general layered upgrade for i with \mathcal{S} there will be a \leq_i -reachable φ -world, $\langle \text{gy}_{\mathcal{S}}^i \rangle \langle \leq_i \rangle \varphi$, if and only if before the operation the current world is not in \mathcal{S} and can $\leq_{\text{def}}^{\mathcal{S}}$ -reach a world not in \mathcal{S} that will satisfy φ after the operation (first disjunct on part 1¹²), or else the current world is in some $\mathcal{S}[k]$ and can $\leq_{\text{def}}^{\mathcal{S}}$ -reach a world also in $\mathcal{S}[k]$ that will satisfy φ after the operation (second disjunct on part 1), or else there is a k such that the current world is not in the sets $\mathcal{S}[1], \dots, \mathcal{S}[k]$ and there is a world in $\mathcal{S}[k]$ that will satisfy φ after the operation (part 2). This is simply the unfolding of the definition of $\leq_{\mathcal{S}}$ (Definition 8), and it emphasises the role played by the formulas characterising each $\chi_{\mathcal{S}[k]}$ and the relational expression characterising $\leq_{\text{def}}^{\mathcal{S}}$.

Observe how the simultaneous upgrade modality $\langle \text{gy}_{\mathcal{S}} \rangle$, briefly sketched below Definition 11, is also axiomatised by the presented system as long as the relational transformer is changed by making the cases for each agent i relative to i 's layered list \mathcal{S}_i (thus removing the cases “for agents different from i ”).

Going back to the example discussed in Section 3, let p_i denote the fact that ‘movie i is most preferred’. Consider the PR model M as given in Example 1, where agent a can be the most reliable agent for himself, and $V(p_i) = \{w_i\}$ for each i . Then one can easily show that $W_M = \llbracket \langle \leq_b \rangle p_2 \wedge \langle \text{gy}_{\mathcal{S}}^b \rangle \langle \leq_b \rangle p_4 \rrbracket^M$.

4 Conclusions and Further Work

The paper has introduced a general preference upgrade operation subsuming several reasonable upgrade policies, providing also a modality to describe its effects as well as its complete axiomatisation. As the motivation for this work comes from the modelling of the process of deliberation, the next step in this research project is the characterisation not only of those situations in which the repetitive application of (instances of) the defined operation leads to agents having the same preferences (preference unanimity), but also of those situations in which further applications of the operation do not make any difference (preference stability). Also interesting is an in-depth exploration of the power of such operation as well as an extensive and formal comparison with related frameworks. Finally, it would be meaningful to get a formal framework for decision making processes that combine the methods of deliberation and aggregation.

¹² Recall that $?(\varphi, \psi)$ describes the relation $\llbracket \varphi \rrbracket^M \times \llbracket \psi \rrbracket^M$. Hence, $\leq \cap ?(\psi, \psi)$ describes the restriction of $\lceil \leq \rceil^M$ to the set of worlds satisfying ψ .

References

1. Dietrich, F., List, C.: Where do preferences come from? *International Journal of Game Theory* **42** (2013) 613–637
2. Dryzek, J.S., List, C.: Social choice theory and deliberative democracy: A reconciliation. *British Journal of Political Science* **33** (2003) 1–28
3. Ghosh, S., Velázquez-Quesada, F.R.: Agreeing to agree: Reaching unanimity via preference dynamics based on reliable agents. In Bordini, R., Elkind, E., Weiss, G., Yolum, P., eds.: *AAMAS 2015*. (2015) 1491–1499
4. Burgess, J.P.: Basic tense logic. In Gabbay, D., Guenther, F., eds.: *Handbook of Philosophical Logic*. Volume II. Reidel (1984) 89–133
5. Goldblatt, R.: *Logics of Time and Computation*. 2nd edn. Number 7 in *CSLI Lecture Notes*. Center for the Study of Language and Information, Stanford, CA (1992)
6. Gargov, G., Passy, S.: A note on boolean modal logic. In Petkov, P.P., ed.: *Mathematical Logic*. Springer US (1990) 299–309
7. Harel, D., Kozen, D., Tiuryn, J.: *Dynamic Logic*. MIT Press, Cambridge, MA (2000)
8. Rott, H.: Shifting priorities: Simple representations for 27 iterated theory change operators. In Lagerlund, H., Lindström, S., Sliwinski, R., eds.: *Modality Matters: Twenty-Five Essays in Honour of Krister Segerberg*. Number 53 in *Uppsala Philosophical Studies*. University of Uppsala, Uppsala (2006) 359–384
9. van Benthem, J.: Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics* **17** (2007) 129–155
10. van Ditmarsch, H., van der Hoek, W., Kooi, B.: *Dynamic Epistemic Logic*. Number 337 in *Synthese Library*. Springer (2008)
11. van Benthem, J.: *Logical Dynamics of Information and Interaction*. Cambridge University Press (2011)
12. van Benthem, J., Liu, F.: Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logics* **17** (2007) 157–182
13. Grüne-Yanoff, T., Hansson, S.O., eds.: *Preference Change*. Volume 42 of *Theory and Decision Library*. Springer (2009)
14. van Eijck, J., Wang, Y.: Propositional dynamic logic as a logic of belief revision. In Hodges, W., de Queiroz, R.J.G.B., eds.: *WoLLIC*. Volume 5110 of *Lecture Notes in Computer Science*. Springer (2008) 136–148
15. Baltag, A., Smets, S.: A qualitative theory of dynamic interactive belief revision. In Bonanno, G., van der Hoek, W., Wooldridge, M., eds.: *Logic and the Foundations of Game and Decision Theory (LOFT7)*. Volume 3 of *Texts in Logic and Games*. Amsterdam University Press, Amsterdam, The Netherlands (2008) 13–60
16. Ghosh, S., de Jongh, D.: Comparing strengths of beliefs explicitly. *Logic Journal of the IGPL* **21** (2013) 488–514
17. Wang, Y., Cao, Q.: On axiomatizations of public announcement logic. *Synthese* **190** (2013) 103–134
18. van Benthem, J., van Eijck, J., Kooi, B.: Logics of communication and change. *Information and Computation* **204** (2006) 1620–1662
19. Baltag, A., Moss, L.S., Solecki, S.: The logic of public announcements and common knowledge and private suspicions. In Gilboa, I., ed.: *Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-98)*, Evanston, IL, USA, July 22–24, 1998, Morgan Kaufmann (1998) 43–56