



Belief extrapolation (or how to reason about observations and unpredicted change) [☆]

Florence Dupin de Saint-Cyr ^{a,*}, Jérôme Lang ^b

^a IRIT, Université Paul Sabatier, 31062 Toulouse Cedex 9, France

^b LAMSADE, Université Paris-Dauphine, 75775 Paris Cedex 16, France

ARTICLE INFO

Article history:

Received 8 September 2009

Received in revised form 31 October 2010

Accepted 31 October 2010

Available online 3 November 2010

Keywords:

Belief change

Belief revision

Reasoning about change

Temporal reasoning

ABSTRACT

We give a logical framework for reasoning with observations at different time points. We call belief extrapolation the process of completing initial belief sets stemming from observations by assuming minimal change. We give a general semantics and we propose several extrapolation operators. We study some key properties verified by these operators and we address computational issues. We study in detail the position of belief extrapolation with respect to revision and update: in particular, belief extrapolation is shown to be a specific form of time-stamped belief revision. Several related lines of work are positioned with respect to belief extrapolation.

© 2010 Published by Elsevier B.V.

1. Introduction

Reasoning about change is a major issue in knowledge representation. Several belief change paradigms have been deeply investigated: *belief revision* is usually viewed as changing an agent's initial belief state to take into account a new piece of information referring to the same situation; more generally, *belief merging* consists in integrating several belief states about the same situation into a single belief state; *belief update* modifies a belief state according to a possible change in the world corresponding to an action effect expressed by a propositional formula. A closely related area of research is *reasoning about action*, where *progression* (consisting in determining the belief state obtained after an action is performed, given the initial belief state and the description of the effects of the action) and conversely *regression*, can be considered as well as classes of belief change operators.

A belief change paradigm that has been considered much less often is the following: given a series of observations obtained at different time points, find out the most plausible events that occurred (and when they occurred). Here we give to this class of problems a deep attention, assuming minimal change (or equivalently, assuming that the world has a tendency towards inertia). We call this belief change paradigm *belief extrapolation*, because it amounts to complete initial beliefs at different time points using some assumptions about how the world evolves.¹

Since belief extrapolation deals with beliefs at different time points, it is worth saying now why it differs from belief update (this will be made more precise in Section 7). Let us discuss it first in the specific case where there are only two time

[☆] A preliminary version of this article appeared in the *Proceedings of the Eighth International Conference on Principles of Knowledge Representation and Reasoning (KR2002)*. This submission includes significant revisions and expansions.

* Corresponding author. Tel.: +33 5 61 55 74 50; fax: +33 5 61 55 62 39.

E-mail addresses: bannay@irit.fr (F. Dupin de Saint-Cyr), lang@lamsade.dauphine.fr (J. Lang).

¹ Since we are interested in such a completion not only on future time points, but more generally on any time point in the history of observations, *belief interpolation* could have been a better name. However, since the conference version of this article [13] makes use of the name *extrapolation*, which has been reused later by several authors, we prefer to minimize change and ambiguity, and keep this name.

points 1 and 2. In this case, extrapolation has as input two formulae (corresponding to the observations at time points 1 and 2), and belief update also has as input two formulae attached to time points 1 and 2. However, the formula corresponding to 2 has a very different interpretation in update and extrapolation.

- *extrapolation*: some time has passed between the initial time point 1 when the initial belief state K was known to hold and a later time point 2 when φ is *observed* (by means of a sensor, a request to a database, a notification by another agent, etc.);
- *update*: φ is the *expected consequence* of an action or an event that occurred between 1 and 2.

Failing to distinguish these types of information may lead to counterintuitive results, as seen on the following example:

Example 1. We consider a system to diagnose, whose components may fail independently (there is no prior causal link between the failure of a component and the failure of another one). The system has a tendency towards inertia: by default, the mode of a given component – working or faulty) persists (in other words, changes of mode are exceptional).

Scenario 1: at $t = 1$, we know that exactly one of components a and b is faulty, and c works correctly: $K = (ok_a \oplus ok_b) \wedge ok_c$ where \oplus denotes exclusive or. After waiting for some time, at $t = 2$, some new observations (coming for instance from tests) make us learn that component b is ok but c is not: $\varphi = ok_b \wedge \neg ok_c$.

Scenario 2: at $t = 1$, like in Scenario 1, $K = (ok_a \oplus ok_b) \wedge ok_c$. Then the action of replacing b by a brand new component as well as the (unfortunate) action of breaking down component c are performed. The effects of these two independent actions can be expressed by $\varphi = ok_b \wedge \neg ok_c$.

These two scenarios are different and lead to different plausible conclusions about the evolution of the world. In Scenario 1, since component b is OK and changes are exceptional, there are good reasons to believe that b was already OK at time 1, so component a was the faulty one and most expectedly still is; there is no ambiguity about c , which has just been observed faulty. Therefore, the belief state at time 2 should be $\neg ok_a \wedge ok_b \wedge \neg ok_c$. In Scenario 2, the effect of replacing b does not bring any new information on its former state, nor, *a fortiori*, on the former state of a ; the belief about a should thus remain unchanged and the belief state at $t = 2$ should be $ok_b \wedge \neg ok_c$.

Belief update may be the right thing to do for Scenario 2, but not for Scenario 1. This inadequacy of belief update to handle Scenario 1 is due to the most “typical” property of belief update, namely U8 (see Section 7) which says that models of K should be updated independently; as a consequence, usual belief update operators satisfy $K \diamond \varphi = ok_b \wedge \neg ok_c$; this is the expected result for Scenario 2 but not for Scenario 1, where we expect $\neg ok_a \wedge ok_b \wedge \neg ok_c$.

Scenario 2 suggests that belief update works when *the input is the expected result (or projection) of an action (or an event) whose occurrence is known to the agent*. Now, how should we change a belief base to take account of observations occurring at different time points, as in Scenario 1? This question leads to defining a different operator that we call *belief extrapolation*. This operator takes a sequence of observations assumed to hold for certain and projects these observations forwards and backwards, assuming that fluents tend to persist throughout time. The basic assumptions for extrapolation operators are the following:

1. *the agent observes some properties of the world at different time points*, but does not have the ability of performing actions; furthermore, if exogenous events occur, they are perceived by the agent through the observations of their effects only (for instance, the occurrence of the event that it rained last night is perceived by actually *having seen* the rain last night, or by seeing the wet ground this morning). This is why changes can be qualified as *unpredicted*.
2. *the system is inertial*: by default, it remains in a static state. This assumption justifies the use of a change minimization policy.

Our main contribution is a thorough investigation of this extrapolation process. Although such processes had been considered in the past (see Section 8), they had never received such a deep attention nor such a systematic study in a belief change perspective. Our contributions are both foundational and computational. On the foundational side, we define a very general and structured class of extrapolation operators. We establish a very precise connection between extrapolation and belief revision, showing that extrapolation can be seen as a particular case of revision over a time-stamped language (Proposition 3). This leads us to give an axiomatic framework for extrapolation: Propositions 4 and 5 are representation theorems inherited from representation theorems for belief revision, whereas Propositions 6, 8 and Lemma 3 give sufficient (and sometimes necessary) conditions for some specific temporal properties (reversibility, Markovianity, independence of empty observations) to be satisfied. We also give an impossibility result (Proposition 16) showing that an extrapolation operator cannot be an update operator. On the computational side, we identify the computational complexity of belief extrapolation (Proposition 12). Then we give a method for computing extrapolated beliefs, for a specific (yet representative enough) operator (Proposition 15).

The rest of the paper is organized as follows. In Section 2 we give the formal definition of belief extrapolation. In Section 3 we propose and discuss several classes of extrapolation operators. In Section 4 we show that belief extrapolation can be seen as a particular belief revision process, where the beliefs about the persistence of fluents are revised by time-stamped observations. In Section 5 we show that extrapolation, however, goes beyond revision, since it enjoys some specific

temporal properties. In Section 6 we investigate computational issues, and propose a practical method for computing extrapolated beliefs for one of the simplest extrapolation operators considered in Section 3. In Section 7, we study the connections between extrapolation and update; we show that a belief extrapolation operator cannot be a belief update operator unless very strong assumptions are made. Connections with several other works are investigated in Section 8, especially dynamic diagnosis, and Section 9 concludes the paper.

2. Belief extrapolation: definitions and general properties

2.1. Temporal formulae, scenarios and trajectories

Let \mathcal{L} be a propositional language built from a finite set of propositional variables (or fluents) \mathcal{V} , the usual connectives, and the Boolean constants \top (tautology) and \perp (contradiction). $\mathcal{M} = 2^{\mathcal{V}}$ is the set of interpretations, or states, for \mathcal{V} . Formulae of \mathcal{L} are denoted by small Greek letters (φ, ψ etc.) and interpretations are denoted by m, m' etc.

An interpretation m is usually denoted by listing the literals it satisfies; for instance, $a-b-c$ denotes the interpretation assigning a to true and b and c to false. If $\varphi \in \mathcal{L}$ then $\text{Mod}(\varphi)$ is the set of all models of φ . \models denotes classical logical consequence. If $X \subseteq \mathcal{M}$ then $\text{form}(X)$ is the propositional formula, unique up to logical equivalence, such that $\text{Mod}(\text{form}(X)) = X$.

A literal is a variable or its negation; if $l = v$ (resp. $\neg v$) then $\neg l$ is $\neg v$ (resp. v). The set of literals associated with \mathcal{V} is $\text{LIT} = \{v, \neg v \mid v \in \mathcal{V}\}$. A formula φ is said to be *complete* if it has only one model, or equivalently, if it is equivalent to a maximal consistent conjunction of literals.

Definition 1 (*Temporal formulae*). Let N be a positive integer (representing the number of time points considered). A *time-stamped propositional variable* is a propositional variable indexed by a time point. If $v \in \mathcal{V}$ and $t \in \llbracket 1, N \rrbracket$ then the intuitive meaning of $v_{(t)}$ is that $v_{(t)}$ holds if and only if v holds at time t . $\mathcal{V}_{(N)} = \{v_{(t)} \mid v \in \mathcal{V}, 1 \leq t \leq N\}$ denotes the set of all time-stamped propositional variables (with respect to \mathcal{V} and N). $\mathcal{L}_{(N)}$ is the language generated from $\mathcal{V}_{(N)}$, the time-stamped Boolean constants $\top_{(t)}$ and $\perp_{(t)}$ ($1 \leq t \leq N$) and the usual connectives.

A formula of $\mathcal{L}_{(N)}$ is called a *temporal formula*. Temporal formulae are denoted by capital Greek letters (Φ, Ψ etc.) For any formula φ of \mathcal{L} and any time point $t \in \llbracket 1, N \rrbracket$, $\varphi_{(t)}$ denotes the temporal formula obtained by time-stamping each variable v appearing in φ by t , i.e., replacing v by $v_{(t)}$. $\varphi_{(t)}$ is called a t -formula.

For instance, $((a \wedge \neg b) \vee c)_{(1)} = (a_{(1)} \wedge \neg b_{(1)}) \vee c_{(1)}$.

Definition 2 (*Scenarios*). A *scenario* is a conjunction of t -formulae for $t \in \llbracket 1, N \rrbracket$, i.e., a temporal formula Σ of the form $\varphi_{(1)}^1 \wedge \dots \wedge \varphi_{(N)}^N$, where $\varphi^1, \dots, \varphi^N$ are formulae of \mathcal{L} . To simplify notations, a scenario Σ is written as a finite sequence of propositional formulae: $\Sigma = \langle \varphi^1, \dots, \varphi^N \rangle$, where $\Sigma(i) = \varphi^i$. $\mathcal{S}_{(N)}$ denotes the set of all scenarios.²

The intuitive motivation for introducing scenarios is that in practice, temporal formulae may often result from a series of observations at different (but precisely located) time points. However, considering scenarios *only* is not sufficient since it does not allow for expressing dependencies between fluents at different time points, neither observations whose temporal location is imprecise.

Example 2. Let $\mathcal{V} = \{a, b\}$ and $N = 3$. $a_{(1)} \wedge (a_{(2)} \vee \neg b_{(2)}) \wedge \top_{(3)}$ is a scenario, which we denote more simply by $\langle a, a \vee \neg b, \top \rangle$. $a_{(1)} \vee b_{(2)}$ is a temporal formula (but not a scenario).

Definition 3 (*Trajectories*). $\text{TRAJ}_{(N)} = 2^{\mathcal{V}_{(N)}}$ denotes the set of all interpretations for $\mathcal{V}_{(N)}$, called *temporal interpretations*, or *trajectories*. For the sake of notation, a trajectory τ will be denoted as a sequence $\tau = \langle \tau(1), \dots, \tau(N) \rangle$ of interpretations in \mathcal{M} . Since a trajectory τ is nothing but an interpretation for the vocabulary $\mathcal{V}_{(N)}$, and a temporal formula is based on the same vocabulary, the satisfaction of a temporal formula $\Phi \in \mathcal{L}_{(N)}$ by a trajectory $\tau \in \text{TRAJ}_{(N)}$ is defined in the usual way (as in standard propositional logic) and is denoted by $\tau \models \Phi$. $\text{Traj}(\Phi) = \{\tau \in \text{TRAJ}_{(N)} \mid \tau \models \Phi\}$ is the set of trajectories satisfying Φ . A temporal formula Φ is consistent if and only if $\text{Traj}(\Phi) \neq \emptyset$.

A trajectory τ is *static* if and only if $\tau(1) = \dots = \tau(N)$.

If τ is a trajectory and $v \in \mathcal{V}$, we denote by $\tau(t)(v)$ the truth value of v at time t in τ , that is, $\tau(t)(v) = \text{true}$ if $\tau(t) \models v$ and $\tau(t)(v) = \text{false}$ if $\tau(t) \models \neg v$.

Note that if Σ is a scenario, then $\tau \models \Sigma$ if and only if $\forall t \in \llbracket 1, N \rrbracket$ we have $\tau(t) \models \Sigma(t)$. Moreover, a scenario Σ is consistent if and only if for each $i \in \llbracket 1, N \rrbracket$, $\Sigma(i)$ is consistent.

² Formally, we should write $\mathcal{S}_{\mathcal{V}, N}$, since the set of scenarios has been defined for a given set of propositional variables \mathcal{V} and a fixed N . However, for the sake of notation we prefer to omit the subscripts whenever this does not induce any ambiguity.

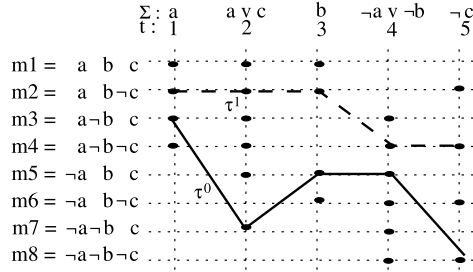


Fig. 1. Two trajectories satisfying $\Sigma = \langle a, a \vee c, b, \neg a \vee \neg b, \neg c \rangle$.

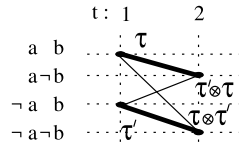


Fig. 2. Trajectories satisfying $\Phi = \langle a(1) \leftrightarrow a(2) \rangle \wedge b(1) \wedge \neg b(2)$ and $S(\Phi)$.

Example 3. Let us consider the scenario $\Sigma = \langle a, a \vee c, b, \neg a \vee \neg b, \neg c \rangle$. The trajectories satisfying Σ are all those connecting the big dots of Fig. 1. Two of them are represented: $\tau^0 = \langle m3, m7, m5, m5, m8 \rangle$ and $\tau^1 = \langle m2, m2, m2, m4, m4 \rangle$.

A temporal formula is not necessarily expressible by a scenario. However, it is possible to define a scenario which is a conjunctive approximation of a temporal formula. This scenario is obtained by projecting the formula at each time point:

Definition 4 (Scenario approximation of a temporal formula). Given a temporal formula $\Psi \in \mathcal{L}_{(N)}$, its scenario approximation is the scenario $S(\Psi) = \langle \Psi^{\downarrow 1}, \dots, \Psi^{\downarrow N} \rangle$, where $\Psi^{\downarrow t}$ is the formula of \mathcal{L} , unique up to logical equivalence such that $\text{Mod}(\Psi^{\downarrow t}) = \{ \tau(t) \mid \tau \in \text{Traj}(\Psi) \}$.

Note that the projection $\Psi^{\downarrow t}$ is the result of forgetting all variables except those concerning time t in Ψ ($\Psi^{\downarrow t} = \text{Forget}(\{v_{(t')} \mid v_{(t')} \in \mathcal{V}_{(N)} \text{ and } t' \neq t\}, \Psi)$) where forgetting a set of variables F in a propositional formula φ is defined inductively as follows [24]:

$$\begin{aligned} \text{Forget}(\{x\}, \varphi) &= \varphi_{x \leftarrow \perp} \vee \varphi_{x \leftarrow \top}, \\ \text{Forget}(\{x\} \cup F, \varphi) &= \text{Forget}(\{x\}, \text{Forget}(F, \varphi)). \end{aligned}$$

Note also that $\Psi^{\downarrow t}$ is the strongest formula (unique up to logical equivalence) implied by Ψ in which only variables of $\mathcal{V}_{(t)}$ appear (see [21]).

Intuitively, $S(\Psi)$ is obtained from Ψ by forgetting the dependencies between the pieces of information pertaining to different time points; equivalently, $S(\Psi)$ is the strongest temporal formula such that (a) $\Psi \models S(\Psi)$ and (b) for every $t \in \llbracket 1, N \rrbracket$ and every formula $\varphi \in \mathcal{L}$, $\Psi \models \varphi_{(t)}$ if and only if $S(\Psi) \models \varphi_{(t)}$.

Note that the trajectories satisfying $S(\Psi)$ are obtained by “crossing” the trajectories in $\text{Traj}(\Psi)$. If τ^1, \dots, τ^N are N trajectories, let $\tau^1 \otimes \dots \otimes \tau^N$ be the trajectory τ defined by: $\forall t \in \llbracket 1, N \rrbracket, \tau(t) = \tau^t(t)$, i.e., $\tau(1) = \tau^1(1), \dots, \tau(N) = \tau^N(N)$. We have the following fact (whose proof is a straightforward rewriting of the definition of a scenario approximation):

Observation 1. $\tau \in \text{Traj}(S(\Psi))$ if and only if there exist N trajectories τ^1, \dots, τ^N in $\text{Traj}(\Psi)$ such that $\tau = \tau^1 \otimes \dots \otimes \tau^N$.

Observation 2. S is monotonic with respect to logical consequence.

Proof. Let us consider two formulae $\Psi, \Psi' \in \mathcal{L}_{(N)}$ such that $\Psi \models \Psi'$. Let us show that $S(\Psi) \models S(\Psi')$. Let $\tau \in \text{TRAJ}_N$ such that $\tau \models S(\Psi)$. From Observation 1, we get that $\exists \tau^1, \dots, \tau^N$ in $\text{Traj}(\Psi)$ such that $\tau = \tau^1 \otimes \dots \otimes \tau^N$. Because $\Psi \models \Psi'$, it holds that τ^1, \dots, τ^N belong to $\text{Traj}(\Psi')$. Hence, using again Observation 1, $\tau \models S(\Psi')$. \square

Example 4. Let $N = 2$, $\mathcal{V} = \{a, b\}$ and consider the temporal formula $\Phi = \langle a(1) \leftrightarrow a(2) \rangle \wedge b(1) \wedge \neg b(2)$. Φ is satisfied by two trajectories: $\tau = \langle ab, a \neg b \rangle$ and $\tau' = \langle \neg ab, \neg a \neg b \rangle$, as shown in Fig. 2. The scenario approximation of Φ is the scenario $S(\Phi) = \langle b, \neg b \rangle$. $S(\Phi)$ is (obviously) satisfied by τ and τ' , but also by two more trajectories: $\tau \otimes \tau' = \langle ab, \neg a \neg b \rangle$, and $\tau' \otimes \tau = \langle \neg ab, a \neg b \rangle$. What has been lost when approximating Φ into $S(\Phi)$ is the dependency between $a(1)$ and $a(2)$.

In the rest of the paper we often refer to the *change set* of a trajectory, that is, the set of all pairs consisting of a literal and a time point such that the literal becomes true at this time point:

Definition 5 (*Change set*). The *change set* $Ch(\tau)$ of a trajectory τ is defined by:

$$Ch(\tau) = \{ \langle l, t \rangle \mid l \in LIT, t \in \llbracket 2, N \rrbracket, \tau(t-1) \models \neg l \text{ and } \tau(t) \models l \}.$$

We also use the following notions, given a change set γ :

- $\gamma(t) = \{ l \mid \langle l, t \rangle \in \gamma \}$ is the set of literals changing to true at time point t in the change set γ
- $ChTime(\gamma) = \{ t \mid \gamma(t) \neq \emptyset \}$ is the set of time points where at least one change occurs in γ .
- $\gamma(l) = \{ t \mid \langle l, t \rangle \in \gamma \}$ is the set of time points where l becomes true in γ .
- $ChLit(\gamma) = \{ l \in LIT \mid \gamma(l) \neq \emptyset \}$ is the set of literals that change to true at least once in γ .

Example 3 (*Continued*).

$$Ch(\tau^0) = \{ \langle \neg a, 2 \rangle, \langle b, 3 \rangle, \langle \neg b, 5 \rangle, \langle \neg c, 5 \rangle \} \text{ and } Ch(\tau^1) = \{ \langle \neg b, 4 \rangle \};$$

$$Ch(\tau^0)(5) = \{ \neg b, \neg c \}; \quad ChTime(Ch(\tau^0)) = \{ 2, 3, 5 \};$$

$$Ch(\tau^0)(\neg b) = \{ 5 \}; \quad ChLit(Ch(\tau^0)) = \{ \neg a, b, \neg b, \neg c \}.$$

Note that a trajectory τ can be unambiguously defined by one of its states (for instance, its initial state $\tau(1)$) and its change set.

2.2. Preferred trajectories and extrapolation operators

Given an observation scenario Σ , or more generally a temporal formula Φ , *belief extrapolation* consists in completing it by using persistence assumptions. Such a process is a specific case of *chronicle completion* [27]. The rationale of belief extrapolation is that as long as nothing tells the contrary, fluents do not change. Semantically, extrapolation consists in finding the preferred trajectories satisfying Φ , with respect to some given preference relation between trajectories (this is similar to many approaches to non-monotonic reasoning, where we select preferred models among those that satisfy a formula).

Definition 6 (*Preference relations on trajectories*). A *preference relation* \preceq is a reflexive and transitive relation on $TRAJ_{(N)}$ (not necessarily connected). $\tau \preceq \tau'$ means that τ is at least as preferred³ as τ' . As usual, we write $\tau < \tau'$ (τ is strictly preferred to τ') for ($\tau \preceq \tau'$ and not $\tau' \preceq \tau$), and $\tau \sim \tau'$ for ($\tau \preceq \tau'$ and $\tau' \preceq \tau$).

For any $X \subseteq TRAJ_{(N)}$, a trajectory $\tau \in X$ is *minimal* w.r.t. \preceq in X if and only if there is no τ' in X such that $\tau' < \tau$. The set of all minimal trajectories w.r.t. \preceq in X is denoted by $\min(\preceq, X)$.

Definition 7 (*Inertial and change-based preference relations*). Let \preceq be a preference relation on $TRAJ_{(N)}$.

- \preceq is *inertial* if for any two trajectories $\tau, \tau' \in TRAJ_{(N)}$:
 - (i) if τ and τ' are both static then $\tau \sim \tau'$;
 - (ii) if τ is static and τ' is not then $\tau < \tau'$.
- \preceq is *change-based* if there exists a reflexive and transitive relation \sqsubseteq on $2^{LIT \times \llbracket 2, N \rrbracket}$ such that for any two trajectories τ, τ' we have $\tau \preceq \tau'$ if and only if $Ch(\tau) \sqsubseteq Ch(\tau')$.
- \preceq is *change-monotonic* if it is change-based, and if for any $\tau, \tau' \in TRAJ_{(N)}$, $Ch(\tau) \subseteq Ch(\tau')$ implies $\tau \preceq \tau'$; \preceq is *strictly change-monotonic* if it is change-monotonic and for any $\tau, \tau' \in TRAJ_{(N)}$, $Ch(\tau) \subset Ch(\tau')$ implies $\tau < \tau'$.

Observation 3.

(O1) \preceq is change-based iff for any τ, τ' , $Ch(\tau) = Ch(\tau')$ implies $\tau \sim \tau'$.

(O2) Any strictly change-monotonic preference relation is inertial.

Proof.

O1: We first show that O1 holds. From left to right, this is straightforward. From right to left: assume that for any τ, τ' , $Ch(\tau) = Ch(\tau')$ implies $\tau \sim \tau'$ and define \sqsubseteq by: for any change sets γ_1, γ_2 , $\gamma_1 \sqsubseteq \gamma_2$ if $\tau_1 \preceq \tau_2$, for some τ_1 and τ_2 such

³ Here, preference has to be interpreted in terms of plausibility, and not in decision-theoretic terms: $\tau \preceq \tau'$ means that τ is *at least as plausible* as τ' .

that $Ch(\tau_1) = \gamma_1$ and $Ch(\tau_2) = \gamma_2$ (such trajectories obviously exist). \sqsubseteq is reflexive and transitive because \preceq is, and for any $\tau, \tau', \tau \preceq \tau'$ if and only if $Ch(\tau) \sqsubseteq Ch(\tau')$.

O2: Assume \preceq is strictly change-monotonic. If τ and τ' are static then $Ch(\tau) = Ch(\tau') = \emptyset$, hence from (O1) and the fact that \preceq is change-based we get $\tau \sim \tau'$. Lastly, if τ is static and τ' is not then $Ch(\tau) = \emptyset$ and $Ch(\tau') \neq \emptyset$, therefore $Ch(\tau') \supset Ch(\tau)$ and $\tau < \tau'$. \square

In the rest of the paper, most preference relations are inertial and change-monotonic.⁴ We are now in position to formally define an extrapolation operator.

Definition 8 (Extrapolation operator). Let \preceq be an inertial preference relation on $TRAJ_{(N)}$. The *extrapolation operator* induced by \preceq maps every temporal formula Φ to another temporal formula $E_{\preceq}(\Phi)$, unique up to logical equivalence, which is satisfied exactly by the *preferred* trajectories among those that satisfy Φ . More formally, E_{\preceq} is a mapping from $\mathcal{L}_{(N)}$ to $\mathcal{L}_{(N)}$ such that

$$Traj(E_{\preceq}(\Phi)) = \min(\preceq, Traj(\Phi)).$$

If \preceq is complete then E_{\preceq} is said to be a complete extrapolation operator.

Note that we required \preceq to be inertial. In principle, we could perfectly well define extrapolation operators based on any preference relation, not necessarily satisfying inertia. This would amount to saying that the world is by default not static. In this paper, however, the assumption that the world is static is important. When discussing the various properties of belief extrapolation as well as its connections to revision and update we will sometimes mention whether or not they are preserved if the inertia assumption is not made.

When there is no ambiguity about the preference relation \preceq , E_{\preceq} is written E .

As said before, we give special attention to the case where the input is a *scenario* and we define the operator that associates with each initial scenario an *extrapolated scenario* defined as the scenario projection of its extrapolation. Formally: Ex is the mapping from $\mathcal{S}_{(N)}$ to $\mathcal{S}_{(N)}$ defined by

$$Ex(\Sigma) = S(E(\Sigma)).$$

The following semantical characterization of scenario-scenario extrapolation will be extremely useful:

Proposition 1.

$$Traj(Ex(\Sigma))(t) = \{\tau(t) \mid \tau \in \min(\preceq, Traj(\Sigma))\}.$$

Proof. $Ex(\Sigma) = S(E(\Sigma))$. Therefore, from Observation 1 we get $\tau \models Ex(\Sigma)$ if and only if there exist τ^1, \dots, τ^N in $Traj(E(\Sigma))$ such that $\tau = \tau^1 \otimes \dots \otimes \tau^N$, that is, if and only if there exist τ^1, \dots, τ^N in $\min(\preceq, Traj(\Sigma))$ such that for every t , $\tau(t) = \tau^t(t)$. \square

Corollary 2.1. $Ex(\Sigma) = \langle \sigma(1), \dots, \sigma(N) \rangle$ where for every t , $Mod(\sigma(t)) = \{\tau(t) \mid \tau \in \min(\preceq, Traj(\Sigma))\}$.

This corollary is straightforward, and so is the following:

Observation 4.

1. For any scenario Σ , we have $E(\Sigma) \models Ex(\Sigma) \models \Sigma$.
2. For any temporal formula Φ , we have $E(\Phi) \models S(E(\Phi)) \models S(\Phi)$ and $E(\Phi) \models \Phi \models S(\Phi)$.

Example 5. Let $\mathcal{V} = \{a\}$ and $N = 2$. The four possible trajectories are $\tau^1 = \langle a, a \rangle$, $\tau^2 = \langle a, \neg a \rangle$, $\tau^3 = \langle \neg a, a \rangle$ and $\tau^4 = \langle \neg a, \neg a \rangle$. Consider the preference relation \preceq defined by $\tau^1 \sim \tau^4 < \tau^2 \sim \tau^3$.

- Let $\Sigma = \langle \top, \top \rangle$. We have $Traj(\Sigma) = \{\tau^1, \tau^2, \tau^3, \tau^4\}$ and $\min(\preceq, Traj(\Sigma)) = \{\tau^1, \tau^4\}$, from which we get $E_{\preceq}(\Sigma) = (a_{(1)} \wedge a_{(2)}) \vee (\neg a_{(1)} \wedge \neg a_{(2)})$ and $Ex_{\preceq}(\Sigma) = Ex_{\preceq}(\top_{(1)} \wedge \top_{(2)}) = \langle \top, \top \rangle$.
- Let $\Phi = a_{(1)} \oplus a_{(2)}$. $E(\Phi) = (a_{(1)} \wedge \neg a_{(2)}) \vee (\neg a_{(1)} \wedge a_{(2)})$, and $Mod(S(E(\Phi)))(1) = \{\tau^2(1), \tau^3(1)\} = \{a, \neg a\}$, and similarly for $t = 2$. Finally, $Ex(S(\Phi)) = \langle \top, \top \rangle$.

⁴ Examples of preference relations which are not change-based are relations where the *context* of the change is relevant. For instance, the relation in which a trajectory is preferred if it satisfies $a \wedge b$ as soon as possible, defined as follows, is not change-based:

$$\tau \preceq \tau' \text{ if and only if } \min\{t, \tau(t) \models a \wedge b\} \leq \min\{t, \tau'(t) \models a \wedge b\}.$$

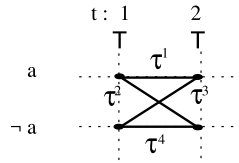


Fig. 3. The four trajectories satisfying $\Sigma = \langle T, T \rangle$.

This example shows that unlike formula–formula extrapolation, scenario–scenario extrapolation generally leads to a loss of information. This is due to the fact that for a given set X of trajectories there is generally no scenario Σ such that $Traj(\Sigma) = X$.

3. Some extrapolation operators

We now give several examples of typical, inertial preference relations and associated extrapolation operators. Although many applications need preference relations that are not change-based, the family of change-based extrapolation operators include a lot of simple and natural operators, and focusing on them is a good start. Since change-based preference relations are induced from a preference relation \sqsubseteq on change sets, therefore it is generally simpler to define them by the corresponding preference relation on change sets.

Recall that the change set of a trajectory τ contains all pairs $\langle f, t \rangle$ such that literal f becomes true at time t in τ . Although there are as many change-based preference relations as preference relations on the set of all possible change sets, there is a reasonable number of prototypical change-based relations obtained by making some natural neutrality assumptions on fluents and on time points. If this neutrality assumption is relaxed, then it is possible to refine the preference relations by considering that changes concerning some particular fluent (or time points) are more important than changes concerning other fluents (or time points). We now give a structured list containing all those natural change-based preference relations (and their associated extrapolation operators) then we discuss the extension of these natural principles to non-neutral preference relations.

3.1. Natural change-based preference relations

To define a natural preference relation on change sets, there are two relevant parameters (see Fig. 4): the part of the change set that suffices to define the preference relation, and the principle used to compare the relevant pieces of information in two change sets (as we shall see, these two parameters are not entirely independent). The relevant information from the change sets that suffices to define the preference relation may concern:

- only the time points where some change occurs;
- only the literals involved in some change;
- only the variables involved in some change;
- the pairs (time point, literal) involved in some change, *i.e.*, the whole change set;
- the pairs (time point, variable) involved in some change.

Finally, a change-based preference relation that we call *basic* needs only the information whether the change set is empty or not. For each of these, we need to distinguish whether several occurrences are counted for once or many (for instance, whether k changes at the same time point should count for k changes or just for one), and, in some cases, whether they should be ranked along with their chronological order of occurrence. To distinguish several occurrences, we make use of multisets. Formally, a multiset μ on a reference set S is a mapping from S to \mathbb{N} , where $\mu(s)$ is the number of occurrences of s in μ . Cardinality and inclusion of multisets are defined as usually $|\mu| = \sum_{s \in S} \mu(s)$ and $\mu \subseteq \mu'$ if $\mu(s) \leq \mu'(s)$ holds for every $s \in S$.

In the following, let γ and γ' be two change sets.

3.1.1. Time points only

Recall that $ChTime(\gamma)$ is the set of time points where at least one change occurs in γ . The neutral two preference relations that can be defined from $ChTime(\gamma)$ make use of set cardinality and set inclusion:

nct $\gamma \sqsubseteq \gamma'$ if $|ChTime(\gamma)| \leq |ChTime(\gamma')|$.

ict $\gamma \sqsubseteq \gamma'$ if $ChTime(\gamma) \subseteq ChTime(\gamma')$.

\preceq_{nct} (*nct* standing for *number of change time points*) prefers trajectories in which changes occur fewer time points, whereas \preceq_{ict} (*ict* standing for *inclusion of change time points*) considers two trajectories as incomparable as soon as each of them contains a time point where a change occurs in it and not in the other one.

Elements	Structure	Comparison principle	Abbreviation
time-points	set	cardinality inclusion	nct ict
	multi-set	cardinality inclusion	nc ncpt
literals (resp. variables)	set	cardinality inclusion	ncl (resp. ncv) icl (resp. icv)
	multi-set	cardinality inclusion	nc (resp. nc) ncpl (resp. ncpv)
	sequence	dominance	cseq (resp. cseqv)
(literal, time point) (resp. (variable, time point))	set	cardinality inclusion	nc (resp. nc) csi (resp. csiv)
boolean (true if there is a change)	singleton	identity	basic

Fig. 4. The 15 natural neutral preference relations.

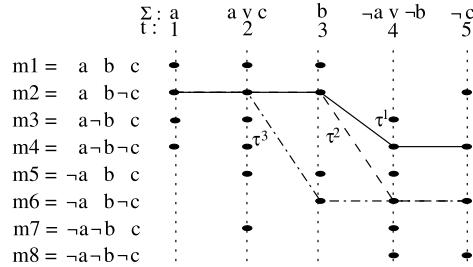


Fig. 5. The 3 \preceq_{nc} -preferred trajectories satisfying $\Sigma = (a, a \vee c, b, -a \vee -b, -c)$.

Example 3 (Continued). The trajectories $\tau^7 = \langle m1, m1, m6, m6, m6 \rangle$ and $\tau^1 = \langle m2, m2, m2, m4, m4 \rangle$ belongs to the set of the 10 trajectories that satisfy $\Sigma = (a, a \vee c, b, -a \vee -b, -c)$ and that are preferred for \preceq_{nct} . All these 10 preferred trajectories have only one instant of change (concerning one or several literals).

There are 26 \preceq_{ict} -preferred trajectories, including the 10 \preceq_{nct} -preferred trajectories and others, namely $\langle m4, m5, m5, m5, m2 \rangle$ for instance.

Now, $ChTime(\gamma)$ does not distinguish time points with a single change from time points with several changes. Define $MChTime(\gamma)$ as the multiset containing, for every time point t , the number of occurrences of changes $\langle f, t \rangle$ in γ : $MChTime(\gamma)$ is a mapping from T to \mathbb{N} , and $MChTime(\gamma)(t)$ is the number of pairs $\langle f, t \rangle$ in $Ch(\tau)$. The natural criteria used above are still relevant:

- nc** $\gamma \sqsubseteq \gamma'$ if $|MChTime(\gamma)| \leq |MChTime(\gamma')|$.
- ncpt** $\gamma \sqsubseteq \gamma'$ if $MChTime(\gamma) \subseteq MChTime(\gamma')$.

\preceq_{nc} simply prefers trajectories with a minimal number of changes, whereas \preceq_{ncpt} prefers trajectories minimizing the number of changes per time point.

Example 3 (Continued). The following 3 trajectories satisfying Σ are preferred for \preceq_{nc} :

- $\tau^1 = \langle m2, m2, m2, m4, m4 \rangle$,
- $\tau^2 = \langle m2, m2, m2, m6, m6 \rangle$,
- $\tau^3 = \langle m2, m2, m6, m6, m6 \rangle$

with the respective change sets $\{\langle -b, 4 \rangle\}$, $\{\langle -a, 4 \rangle\}$ and $\{\langle -a, 3 \rangle\}$ (see Fig. 5). The same trajectories are preferred for \preceq_{nct} together with the trajectory $\langle m1, m5, m5, m5, m6 \rangle$ (with a change set: $\{\langle -a, 2 \rangle, \langle -c, 5 \rangle\}$).

3.1.2. Literals (resp. variables) only

Recall that $ChLit(\gamma)$ is the set of literals that change to true at least once in γ . Let $MChLit(\gamma)$ be the multiset containing, for every literal l , the number of times a change $\langle l, t \rangle$ is present in γ . Then we get the following three natural preference relations:

- ncl** $\gamma \sqsubseteq \gamma'$ if $|ChLit(\gamma)| \leq |ChLit(\gamma')|$.
- icl** $\gamma \sqsubseteq \gamma'$ if $ChLit(\gamma) \subseteq ChLit(\gamma')$.
- ncpl** $\gamma \sqsubseteq \gamma'$ if $MChLit(\gamma) \subseteq MChLit(\gamma')$.

\preceq_{ncl} (resp. \preceq_{icl}) prefers trajectories minimizing the number (resp. the set) of literals that change at least once. \preceq_{ncpl} prefers the trajectories minimizing the number of changes per literal. The fourth natural preference relations that we could define by $|MChLit(\gamma)| \leq |MChLit(\gamma')|$, simply compares trajectories according to the number of changes they contain but it coincides with the already defined preference relation \preceq_{nc} .

Example 3 (Continued). The trajectories τ^1 , τ^2 and τ^3 are the \preceq_{ncl} , \preceq_{icl} and \preceq_{ncpl} preferred trajectories satisfying Σ .

Now, beyond considering sets or multisets of changing literals, we may consider a richer structure, where the information on the *relative order* of changes is kept. The reason for that is that we may want to compare *change sequences* independently of the precise location of changes in time: only the changes involved and the order in which they occurred are relevant. Consider the sequence $SChLit(\gamma)$ obtained from $ChLit(\gamma)$ by clustering the changes that occur at the same time point and ranking these clusters chronologically. More formally, write $ChTime(\gamma) = \{t_1, \dots, t_{|ChTime(\gamma)|}\}$ where $t_1 < \dots < t_{|ChTime(\gamma)|}$, then $SChLit(\gamma)$ is the sequence of sets of literals, of length $|ChTime(\gamma)|$ defined by $SChLit(\gamma)_i = \gamma(t_i)$ for $i \in \llbracket 1, |ChTime(\gamma)| \rrbracket$. For instance, $SChLit(Ch(\tau^0)) = \{\{-a\}, \{b\}, \{-b, -c\}\}$.

cseq $\gamma \sqsubseteq \gamma'$ if there is an injective monotonic mapping F from $\llbracket 1, |SChLit(\gamma)| \rrbracket$ to $\llbracket 1, |SChLit(\gamma')| \rrbracket$, such that: $\forall t \in \llbracket 1, |SChLit(\gamma)| \rrbracket, SChLit(\gamma)(t) \subseteq SChLit(\gamma')(F(t))$

Note that $\tau \sim_{cseq} \tau'$ if and only if the literals which change in τ and in τ' are the same *and strictly in the same order*. We call the preference relation \preceq_{cseq} *change sequence dominance*.

Example 3 (Continued). The three preferred trajectories with respect to \preceq_{cseq} are τ^1 , τ^2 and τ^3 . Notice that $\tau^2 <_{cseq} \tau^4$, that $\tau^4 \sim_{cseq} \tau^5 \sim_{cseq} \tau^6$, and that τ^7 and τ^5 are incomparable, as well as τ^7 and τ^6 : indeed, $Ch(\tau^7) = \{\{-a, 3\}, \{-c, 3\}\}$, i.e., the changes from $-a$ to a and from c to $-c$ occur simultaneously, while for τ^4 , τ^5 and τ^6 , the change from $-a$ to a occurs before the change from c to $-c$.

In some cases, we may want to be indifferent between a change from v to $\neg v$ and a change from $\neg v$ to v , which leads to the following “unsigned” versions of all preference relations based on $ChLit(\gamma)$, $MChLit(\gamma)$ and $SChLit(\gamma)$. Formally, let $\gamma_u = \{\langle v, t \rangle \mid \langle v, t \rangle \in \gamma \text{ or } \langle \neg v, t \rangle \in \gamma\}$, and for every variable v , $\gamma_u(v) = \gamma(v) \cup \gamma(\neg v)$. Then $ChVar(\gamma)$, $MChVar(\gamma)$ and $SChVar(\gamma)$ are defined similarly as $ChLit(\gamma)$, $MChLit(\gamma)$ and $SChLit(\gamma)$, and the corresponding preference relations are defined *mutatis mutandis*. We call these preference relations *unsigned* and denote them by **ncv**, **icv**, **ncpv** and **cseqv**.

Example 3 (Continued). The preferred trajectories are the same for \preceq_{ncv} and \preceq_{icv} ; there are 12, including τ^1 , τ^2 and τ^3 but also trajectories admitting only one changing variable, such as $\langle m_2, m_4, m_2, m_4, m_2 \rangle$, whose change set is $\gamma = \{\langle -b, 2 \rangle, \langle b, 3 \rangle, \langle -b, 4 \rangle, \langle b, 5 \rangle\}$.

3.1.3. Full change set

There are two natural ways of defining preference relations from the whole change set, either by using inclusion or cardinality. If we use inclusion then we obtain \preceq_{csi} (*csi* standing for *change set inclusion*); if we use cardinality, again we obtain \preceq_{nc} , therefore we don't mention it.

csi $\gamma \sqsubseteq \gamma'$ if $\gamma \subseteq \gamma'$.

\preceq_{csi} can also be defined from the *unsigned change set* γ_u :

csiv $\gamma \sqsubseteq \gamma'$ if $\gamma_u \subseteq \gamma'_u$.

Example 3 (Continued). The following 6 trajectories satisfying $\Sigma = \langle a, a \vee c, b, -a \vee -b, -c \rangle$ are preferred for \preceq_{csi} and also for \preceq_{csiv} : τ^1 , τ^2 , τ^3 together with $\tau^4 = \langle m1, m5, m5, m5, m6 \rangle$, $\tau^5 = \langle m1, m5, m5, m6, m6 \rangle$ and $\tau^6 = \langle m1, m5, m6, m6, m6 \rangle$ (whose change sets are, respectively, $\{\langle -a, 2 \rangle, \langle -c, 5 \rangle\}$, $\{\langle -a, 2 \rangle, \langle -c, 4 \rangle\}$ and $\{\langle -a, 2 \rangle, \langle -c, 3 \rangle\}$).

Using cardinality of the unsigned change set would give again \preceq_{nc} , because a positive and negative change of a variable can only occur at different time points, and hence will count for two changes for nc , as it will count for two changes in the unsigned change set.

3.1.4. “Basic” preference relation

In addition to all these, if the relevant information is the binary predicate $Empty(\gamma)$, defined by $Empty(\gamma) = 1$ if $\gamma = \emptyset$ and $Empty(\gamma) = 0$ if $\gamma \neq \emptyset$, then the basic preference relation is defined by:

basic $\gamma \sqsubseteq \gamma'$ if $Empty(\gamma) \geq Empty(\gamma')$.

This degenerate preference relation \preceq_{basic} is inertial and is such that any two non-inertial trajectories are equally preferred.

3.2. Refining natural preference relations by relaxing neutrality

The preference relations defined above can be generalized by giving up the assumption that all time points (respectively literals, variables) are equally important. Levels of importance can be expressed either numerically, associating weights with time points and/or literals or variables, or qualitatively, using priority relations (on time points, literals, etc.).

Because time points are already chronologically rank-ordered, a natural way of breaking neutrality consists in giving priority to later or earlier time points in change sets. For instance, \preceq_{csi} can be refined into a preference relation that prefers trajectories where changes occur as late as possible. This policy is called *chronological minimization* in [31] and is extensively discussed in [27]. It is defined as follows:

chr $\gamma \sqsubseteq \gamma'$ if $\gamma = \gamma'$ or there exists $k \in \llbracket 1, N \rrbracket$ such that $\begin{cases} \gamma(k) \subset \gamma'(k), \\ \forall t < k, \gamma(t) = \gamma'(t). \end{cases}$

Example 3 (Continued). The preferred trajectories for \preceq_{chr} are τ^1 and τ^2 , where the first change occurs at time point 4.

Similar refinements can be done for \preceq_{ict} and \preceq_{ncpt} . However, chronological minimization does not make sense for \preceq_{icl} or \preceq_{ncpl} , since $ChLit(\gamma)$ and $MChLit(\gamma)$ do not convey any information about time points. Lastly, instead of preferring trajectories in which changes occur as late as possible, one can choose to prefer trajectories in which changes occur as early as possible (“anti-chronological minimization”).

Another natural way of breaking neutrality consists in attaching a *penalty* to each possible change. For instance, the following refinement of \preceq_{nc} consists in assuming that the penalty depends only on the literal involved in the change, and that some fluents can change more easily than others. In this case, let $k : LIT \rightarrow \mathbb{N}^*$, where $k(l)$ is the penalty induced by a change from $\neg l$ to l . Note that we do not necessarily have $k(l) = k(\neg l)$, for instance, obviously, $k(alive) < k(\neg alive)$. Now, the cost of a change set γ is defined by:

$$K(\gamma) = \sum_{t=2}^N \sum_{l \in \gamma(t)} k(l).$$

penalty (k) $\gamma \sqsubseteq \gamma'$ if $K(\gamma) \leq K(\gamma')$.

The reason why we impose that changes have a strictly positive penalty is to impose inertia. Note that we recover \preceq_{nc} by letting $k(l) = 1$ for all l .

All the preference relations defined so far in this section consider atomic changes as independent. It may be the case, however, that fluents do not change independently from each other. This is the case when changes are caused by *events*. The following model is taken (and slightly adapted) from [7], where the plausibility of events, as well as their dynamics, are modeled by ordinal conditional functions. Let E be a finite set of events containing the “null” event e_\emptyset . For every event $e \in E$ and every pair of states (m, m') , $\kappa(m' | m, e) \in \mathbb{N} \cup \{\infty\}$ measures the plausibility that m' results when e occurs at m , with the usual convention the higher $\kappa(m' | m, e)$, the more exceptional the transition from m to m' . Moreover, for every state m and event e , $\kappa(e | m) \in \mathbb{N} \cup \{\infty\}$ measures the plausibility that e occurs in m .

Note that since we want to assume that the system is inert, we impose that the two following conditions necessarily hold (see [7] for a discussion, as well as for examples or event models):

1. for any e and m : $\kappa(e | m) = 0$ if and only if $e = e_\emptyset$;
2. for any m, m' : $\kappa(m' | m, e_\emptyset) = 0$ if and only if $m = m'$.

An event model $\mathcal{E} = (E, \kappa)$ consists of E and the two collections of κ -functions described above. Given an event model \mathcal{E} , the plausibility of the transition from m to m' through e is $\kappa_{\mathcal{E}}(e, m, m') = \kappa(e | m) + \kappa(m' | m, e)$, and the plausibility of the transition from m to m' is defined as $\kappa_{\mathcal{E}}(m, m') = \min_{e \in E} \kappa_{\mathcal{E}}(e, m, m')$.

Lastly, the plausibility of a trajectory τ is defined by $\kappa_{\mathcal{E}}(\tau) = \sum_{t=1}^{N-1} \kappa_{\mathcal{E}}(\tau(t), \tau(t+1))$. Note that if τ is static then obviously, $\kappa_{\mathcal{E}}(\tau) = 0$, and if τ is not static then one of the events in the sequence must be different from e_\emptyset and thus $\kappa_{\mathcal{E}}(\tau) > 0$. Hence the conditions imposed on κ well ensure that inertia corresponds to normality, and consequently, the event-based preference relation defined below prefers inert trajectory:

event penalty (\mathcal{E}) $\tau \preceq \tau'$ if $\kappa_{\mathcal{E}}(\tau) \leq \kappa_{\mathcal{E}}(\tau')$.

Note that since the effects of events generally depend on the initial state, knowing $Ch(\tau)$ and $Ch(\tau')$ is generally not enough to decide whether $\tau \preceq_{\mathcal{E}} \tau'$, i.e., event-based preference relations are generally not change-based.

In all, we have obtained 15 “natural” preference relations, assuming neutrality between time points and fluents, and we have given three ways of refining them. Since it would be tedious to study them all, in the rest of the paper we will focus on a few representative, “prototypical” preference relations among these.

4. Extrapolation and revision

We first need to recall some background on belief revision. Rather than the usual presentation of revision by Alchourrón, Gärdenfors and Makinson [1], that views revision as mapping a closed logical theory K and an input formula α to a closed logical theory $K \star \alpha$, we choose here the syntactical presentation of Katsuno and Mendelzon [19], which views revision as mapping a propositional formula φ (representing the initial belief state) and another propositional formula α (the “input”) to a propositional formula $\varphi \star \alpha$ (representing the belief state after revision by α). This simplification is without loss of generality whenever the propositional language is generated by a *finite* set of propositional symbols (which is the case in this paper). Under this assumption, the postulates listed by [1] correspond to the following ones [19]:

- R1** $\varphi \star \alpha \models \alpha$.
- R2** If $\varphi \wedge \alpha$ is satisfiable then $\varphi \star \alpha \equiv \varphi \wedge \alpha$.
- R3** If α is satisfiable then $\varphi \star \alpha$ is satisfiable.
- R4** If $\varphi \equiv \varphi'$ and $\alpha \equiv \beta$ then $\varphi \star \alpha \equiv \varphi' \star \beta$.
- R5** $(\varphi \star \alpha) \wedge \beta \models \varphi \star (\alpha \wedge \beta)$.
- R6** If $(\varphi \star \alpha) \wedge \beta$ is satisfiable then $\varphi \star (\alpha \wedge \beta) \models (\varphi \star \alpha) \wedge \beta$.

Given a propositional language \mathcal{L} , a *revision operator* is a mapping

$$\star: \mathcal{L} \times \mathcal{L} \rightarrow \mathcal{L}$$

satisfying **R1** to **R6**. Katsuno and Mendelzon [19] then establish the following representation theorem (similar to the main representation theorem in [1]):

Proposition 2. (See [19].) \star is a revision operator if and only if there exists an assignment mapping any formula φ to a complete preorder \leq_φ on $\mathcal{M} = 2^{\mathcal{V}}$ for every propositional formula φ , such that

1. the assignment is faithful, i.e.,
 - (1a) if $m \models \varphi$ and $m' \models \varphi$ then $m \leq_\varphi m'$
 - (1b) if $m \models \varphi$ and $m' \models \neg\varphi$ then $m <_\varphi m'$
 (where $<_\varphi$ is defined from \leq_φ by $m <_\varphi m'$ iff $m \leq_\varphi m'$ and not $m' \leq_\varphi m$);
2. $m \models \varphi \star \alpha$ if and only if $m \in \text{Min}(\leq_\varphi, \text{Mod}(\alpha))$.

Belief revision is often thought of as dealing with *static worlds*, therefore with formulae pertaining to the same time point. However, as remarked in [15], “what is important for revision is not that the world is static, but that the propositions used to describe the world are static”. That is, nothing prevents us from considering revision operators in a language generated from a set of time-stamped propositional symbols. Consider then a revision operator \star on $\mathcal{L}_{(N)}$. Because an interpretation of $\mathcal{L}_{(N)}$ corresponds to a trajectory of $\text{TRAJ}_{(N)}$, rewriting the above result for $\mathcal{L}_{(N)}$ leads us to say that it is possible to map faithfully Φ to a complete preorder \leq_Φ on $\text{TRAJ}_{(N)}$ for every temporal formula Φ , such that for every temporal formula Ψ ,

$$(3) \quad \tau \models \Phi \star \Psi \quad \text{if and only if} \quad \tau \in \text{Min}(\leq_\Phi, \text{Traj}(\Psi)).$$

Let us compare (3) with the definition of extrapolation: $\text{Mod}(E_{\preceq}(\Psi)) = \{\tau \mid \tau \in \text{Min}(\preceq, \text{Traj}(\Psi))\}$, or, equivalently,

$$(4) \quad \tau \models E_{\preceq}(\Psi) \quad \text{if and only if} \quad \tau \in \text{Min}(\preceq, \text{Traj}(\Psi)), \quad \text{where } \preceq \text{ is inertial.}$$

Thus, extrapolating a temporal formula Ψ is equivalent to revise *some initial belief state* (that we shall call Φ) by Ψ . What does this initial belief state correspond to? Φ should be such that any given complete preorder assigned faithfully to Φ , \leq_Φ , is inertial. This means that any two static trajectories should be indifferent with respect to \leq_Φ and that if τ is static and τ' is not, then $\tau <_\Phi \tau'$ must hold. Hence, the models of Φ are the *static trajectories*. Let $\text{PERS} = \bigwedge_{v \in V} \bigwedge_{t=1}^{N-1} v_t \leftrightarrow v_{t+1}$ be the temporal formula having for models the set of static trajectories (PERS is true if and only if no change occurs between time 1 and time N). We have just shown above that revising PERS by Ψ is equivalent to extrapolating Ψ . This leads to the following result:

Proposition 3. For every extrapolation operator E_{\preceq} , there exists a revision operator \star on $\mathcal{L}_{(N)}$ such that for every $\Phi \in \mathcal{L}_{(N)}$, $E_{\preceq}(\Phi) \equiv \text{PERS} \star \Phi$.

For any revision operator \star there exists a unique extrapolation operator E such that for every $\Phi \in \mathcal{L}_{(N)}$, $E_{\preceq}(\Phi) \equiv \text{PERS} \star \Phi$.⁵

⁵ Note that if we do not require \preceq to be inertial, a more general version of Proposition 3 holds, where PERS is replaced by a temporal formula Σ such that $\text{Traj}(\Sigma) = \text{Min}(\preceq, \text{TRAJ}_{(N)})$.

Proof. The proof follows easily from the discussion above. Let \preceq be an inertial preference relation on trajectories and E_{\preceq} the associated extrapolation operator. Take any assignment $\Psi \mapsto \leq_{\Psi}$ such that for any $\Psi \equiv PERS$, $\leq_{\Psi} = \preceq$ and for any $\Psi \neq PERS$, \leq_{Ψ} is any ordering satisfying the faithfulness conditions (1) and (2). This assignment is faithful⁶: for $\Psi \neq PERS$ this is obvious from the construction, and for $\Psi \equiv PERS$, (1a) is satisfied because for every two static trajectories τ, τ' we have $\tau \models PERS$ and $\tau' \models PERS$, therefore $\tau \sim \tau'$; (1b) is satisfied because if $\tau \models PERS$ and $\tau' \models \neg PERS$ we have $\tau < \tau'$ (because \preceq is inertial). Finally, define \star by (3). we have $\tau \models PERS \star \Phi$ if and only if $\tau \in \text{Min}(\leq_{PERS}, \text{Traj}(\Phi))$, that is, if and only if $\tau \models E_{\preceq}(\Phi)$.

Conversely, a revision operator \star on $\mathcal{L}_{(N)}$ induces a unique extrapolation operator E_{\star} , defined by $E_{\star} = E_{\leq_{PERS}}$, where \leq_{PERS} is the faithful preorder associated with \star and the initial belief state $PERS$. \square

Therefore, *belief extrapolation is a specific instance of belief revision (on a time-stamped language)*: extrapolation amounts to revising the prior belief that all fluents persist throughout time by the observations. This in turn allows us for deriving easily a representation theorem for extrapolation. We need first to rewrite the Katsuno–Mendelzon postulates in the language of extrapolation:

Let E be an extrapolation operator and Φ a temporal formula.

- E1** $E(\Phi) \models \Phi$.
- E2** If $PERS \wedge \Phi$ is satisfiable then $E(\Phi) \equiv PERS \wedge \Phi$.
- E3** If Φ is satisfiable then $E(\Phi)$ is satisfiable.
- E4** If $\Phi \equiv \Phi'$ then $E(\Phi) \equiv E(\Phi')$.
- E5** $E(\Phi) \wedge \Phi' \models E(\Phi \wedge \Phi')$.
- E6** If $E(\Phi) \wedge \Phi'$ is consistent then $E(\Phi \wedge \Phi') \models E(\Phi) \wedge \Phi'$.

E1 expresses that observations are reliable, therefore extrapolation may only *complete* the observations, without questioning their truth. **E3** expresses that any consistent formula can be extrapolated. **E4** is a syntax-independence property. **E5** and **E6** together are equivalent to: if $E(\Phi) \wedge \Phi'$ is consistent then $E(\Phi \wedge \Phi')$ is equivalent to $E(\Phi) \wedge \Phi'$: when there are some trajectories satisfying both $E(\Phi)$ and Φ' then the trajectories satisfying $E(\Phi \wedge \Phi')$ are obtained by retaining in the trajectories satisfying $E(\Phi)$ those that satisfy Φ' . Finally, **E2** is an inertia assumption: if it can be consistently assumed that no change occurred, then we jump to this conclusion. Note that when Φ is a scenario Σ , (**E2**) comes down to

E2S If $\bigwedge_{t=1}^N \Sigma(t)$ is satisfiable then for every $t \leq N$, $E(\Sigma)(t) \equiv \bigwedge_{t=1}^N \Sigma(t)$.

An immediate – but important – property deriving from **E1**, **E4** and **E5** is that extrapolation is idempotent:

Observation 5. **E1**, **E4** and **E5** imply $E(E(\Phi)) \equiv E(\Phi)$.

Proof. **E1** implies that $E(E(\Phi)) \models E(\Phi)$. Take $\Phi' = E(\Phi)$ in **E5**, we get (a) $E(\Phi) \wedge E(\Phi) \models E(\Phi \wedge E(\Phi))$. Since, from **E1**, we have $E(\Phi) \models \Phi$, it means that (b) $\Phi \wedge E(\Phi) \equiv E(\Phi)$. From (a) and (b) and **E4**, we get $E(\Phi) \models E(E(\Phi))$. \square

We said above that these postulates **E1–E6** correspond to the postulates **R1–R6**. This is expressed more formally as follows:

Lemma 1. Let $E : \mathcal{L}_{(N)} \rightarrow \mathcal{L}_{(N)}$ be a complete extrapolation operator, and \star any revision operator such that E is induced by \star . Then, for any $i \in \{1, \dots, 6\}$, \star satisfies **Ri** if and only if E satisfies **Ei**.

Proof. From Proposition 3, given a revision operator \star there is only one extrapolation operator E induced by \star . Since \star is a revision operator, it satisfies postulates **R1–R6** which are true for every formulae $\varphi, \varphi', \alpha$ and β , taking $PERS$ for φ and for φ' , and taking Φ and Φ' for α and β respectively, **R1–R6** translate immediately into **E1–E6**. \square

Now, the representation theorem (for complete extrapolation operators):

Proposition 4. $E : \mathcal{L}_{(N)} \rightarrow \mathcal{L}_{(N)}$ is a complete extrapolation operator if and only if it satisfies **E1** to **E6**.

Proof. Follows directly from Proposition 3, Lemma 1 and Proposition 2. \square

⁶ Note that a time-stamped propositional language is a standard propositional language, therefore the definition of a faithful assignment does not need to be redefined.

The previous representation theorem characterizes extrapolation operators that are defined on complete inertial orderings, but intuitively, it is natural to expect such orderings to be incomplete, allowing two trajectories to be incomparable. Incomplete extrapolation operators can also be characterized by a set of postulates, as follows:

Proposition 5. *Let us consider the two following postulates:*

E7 If $E(\Phi) \models \Phi'$ and $E(\Phi') \models \Phi$ then $E(\Phi) \equiv E(\Phi')$.

E8 $E(\Phi) \wedge E(\Phi') \models E(\Phi \wedge \Phi')$.

E is a partial extrapolation operator if and only if it satisfies E1–E5, E7, E8.

Proof. The proof is similar to the proof for complete operators (Proposition 4), replacing Proposition 2 by Theorem 5.2 in [19]. \square

Now that we know that, technically speaking, an extrapolation operator is a revision operator on a time-stamped language where the initial belief state is *fixed* (to *PERS*), the question is, *is it nothing more than that?* The answer is both positive and negative. Positive, because indeed, any extrapolation operator can be seen as a revision operator. Negative, because the temporal structure makes extrapolation a specific class of belief change operators with its specific properties, that makes it worth studying on its own: in other terms, extrapolation is *much more specific* than revision, and focusing on this specificity is worth investigating. In Section 5 we will study a few specific temporal properties of extrapolation, which would be meaningless in a standard, atemporal belief revision framework. The discussion on the specificity of extrapolation with respect to plain belief revision will be continued in the conclusion.

We end this section by briefly discussing the possible connections between extrapolation and *iterated belief revision*. As there are quite many works on iterated revision, it would be tedious to compare extrapolation with them all. We choose to compare it with one specific approach, namely Lehmann's iterated revision [22], because it considers sequences of consistent formulae,⁷ therefore the input is, formally speaking, the same as in scenario extrapolation. In [22], an iterated revision function maps any sequence of formulae $\sigma = \langle \varphi_1, \dots, \varphi_n \rangle$ to a belief state $[\sigma]$ resulting from the sequence σ of individual revisions. Since [22] is only concerned with the final result of the iterated revision process, what we can compare given a scenario Σ is the belief resulting from Σ , denoted by $Traj(\Sigma)$, with the projection of the extrapolation at the final time point $Ex(\Sigma)^{\downarrow N}$ (where $N = |\Sigma|$).⁸ The difference between extrapolation and iterated revision is clear when considering the following example: let $\Sigma = \langle a \rightarrow b, a, \neg a \rangle$; any “reasonable” extrapolation operator (including all extrapolation operators defined in Section 3) satisfies $Ex(\Sigma) = \langle a \wedge b, a \wedge b, \neg a \wedge b \rangle$ (the change from a to $\neg a$ between 2 and 3 being certain, the preferred trajectory is the one containing no other changes). Now, Lehmann's iterated revision, and also most iterated revision operators defined on epistemic states (e.g., [4,9]) give $[a \rightarrow b, a, \neg a] = \neg a$. The reason for this difference is that iterated revision is concerned with pieces of information concerning a *static world*; what evolves is the agent's belief state, not the state of the world. Therefore, once the new information $\neg a$ has “cancelled” the preceding one, the reasons to believe in b have disappeared. This strong “directivity” of time in iterated revision contrasts with extrapolation, where past and future can often be interchanged (as soon as the property (R) is satisfied, as we shall see in Section 5).

5. Temporal properties of extrapolation

We now consider *temporal* properties of some extrapolation operators, *i.e.*, properties that explicitly refer to the flow of time. The first and perhaps most important of these temporal properties is *inertia*, which has already been discussed, and which is, by definition, satisfied by all extrapolation operators. We now consider properties that are not always satisfied. For each of these, we discuss the relevance and desirability of the property and we list, among the “typical” extrapolation operators, which ones satisfy it.

5.1. Reversibility

Reversibility expresses that “forward” persistence (inferring beliefs from the past to the future) and “backward” persistence (inferring beliefs from the future to the past) are symmetric. Even if reversibility is not necessarily a desirable property, it is however interesting to see which operators satisfy it and which ones do not, since this property tells a lot about the nature of extrapolation operators. In particular, as we shall see soon, reversibility holds for many of the “typical”

⁷ Other approaches to iterated belief revision act on epistemic states rather than on sequences of formulae and the comparison with extrapolation is more complex to establish.

⁸ Among the postulates for iterated revision [22], it can be shown that the operator $\Sigma \mapsto Ex(\Sigma)^{\downarrow N}$, where Ex is a scenario extrapolation operator, satisfies I1, I2, I3, I5 and I6, and does not satisfy I4 nor I7 (we omit both the postulates and the proofs, since this point is of minor importance). Thus, Lehmann is right when saying that for a belief change operator on a changing situation, “one would probably accept all the postulates [...] except I4 and I7” [22].

extrapolation operators, which sheds more light on how extrapolation departs from (iterated) revision and update, which obviously don't satisfy it.

We first define the *reversal* of a trajectory and of a temporal formula. We recall that N is the number of time points considered.

Definition 9 (*Reversal of trajectories and formulae*). For any trajectory τ , the reversal of τ is defined by

$$\text{reverse}(\tau)(t) = \tau(N - t + 1).$$

For any temporal formula Φ , the reversal of Φ is:

$$\text{Reverse}(\Phi) = \Phi'$$

such that Φ' is the formula Φ in which each occurrence of a variable $x_{(t)}$ is replaced by $x_{(N-t+1)}$.

Remark that in the particular case where Φ is a scenario $\Sigma = \langle \Sigma(1), \dots, \Sigma(N) \rangle$, the definition simplifies to $\text{Reverse}(\Sigma) = \langle \Sigma(N), \dots, \Sigma(1) \rangle$. Next, we introduce the *reversibility* property:

Definition 10 (*Reversibility*). An extrapolation operator satisfies reversibility (R) if and only if for every temporal formula $\Phi \in \mathcal{L}_{(N)}$, we have $\text{Reverse}(E(\Phi)) = E(\text{Reverse}(\Phi))$.

We have the following semantical characterization of reversibility. A preference relation \preceq is said to be *reversible* if and only if for every $\tau, \tau' \in \text{TRAJ}_{(N)}$, $\tau < \tau'$ holds if and only if $\text{reverse}(\tau) < \text{reverse}(\tau')$.

Proposition 6. E_{\preceq} satisfies (R) if and only if \preceq is reversible.

Proof. Let \preceq be a reversible preference relation over $\text{TRAJ}_{(N)}$, and consider the extrapolation operator E_{\preceq} . We have the following chain of equivalences:

$$\begin{aligned} \tau &\models \text{Reverse}(E_{\preceq}(\Phi)) \\ \Leftrightarrow \forall t \in [1..N], \tau(t) &\models \text{Reverse}(E_{\preceq}(\Phi))(t) \\ \Leftrightarrow \forall t \in [1..N], \tau(N - t + 1) &\models \text{Reverse}(E_{\preceq}(\Phi))(N - t + 1) \\ \Leftrightarrow \text{reverse}(\tau) &\models E_{\preceq}(\Phi) \\ \Leftrightarrow \text{reverse}(\tau) &\models \Phi \text{ and } \nexists \tau' \models \Phi \text{ s.t. } \tau' < \text{reverse}(\tau) \\ \Leftrightarrow \tau &\models \text{Reverse}(\Phi) \text{ and } \nexists \tau'' \models \text{Reverse}(\Phi) \text{ s.t. } \text{reverse}(\tau'') < \text{reverse}(\tau) \\ \Leftrightarrow \tau &\models \text{Reverse}(\Phi) \text{ and } \nexists \tau'' \models \text{Reverse}(\Phi) \text{ s.t. } \tau'' < \tau \quad (\text{reversibility of } \preceq) \\ \Leftrightarrow \tau &\in \min(\preceq, \text{Traj}(\text{Reverse}(\Phi))) \\ \Leftrightarrow \tau &\models E_{\preceq}(\text{Reverse}(\Phi)). \end{aligned}$$

Hence, E_{\preceq} satisfies (R).

Conversely, assume \preceq is not reversible. Then there exist two trajectories τ, τ' such that $\tau < \tau'$ and $\text{reverse}(\tau) \not< \text{reverse}(\tau')$. Let Φ be the temporal formula (unique up to logical equivalence) such that $\text{Traj}(\Phi) = \{\tau, \tau'\}$. Because $\tau < \tau'$, we have $\tau' \not\models E_{\preceq}(\Phi)$. Now, $\text{Traj}(\text{Reverse}(\Phi)) = \{\text{reverse}(\tau), \text{reverse}(\tau')\}$, and since $\text{reverse}(\tau) \not< \text{reverse}(\tau')$, we have $\text{reverse}(\tau') \models E_{\preceq}(\text{Reverse}(\Phi))$, that is, $\tau' \models \text{Reverse}(E_{\preceq}(\text{Reverse}(\Phi)))$. Therefore, $E_{\preceq}(\Phi) \not\models \text{Reverse}(E_{\preceq}(\text{Reverse}(\Phi)))$, that is, $\text{Reverse}(E_{\preceq}(\Phi)) \not\models E_{\preceq}(\text{Reverse}(\Phi))$, hence E_{\preceq} does not satisfy (R). \square

Reversibility is strongly linked to the equal plausibility of a change from a to $\neg a$ and *vice versa*, as it can be seen on the following lemma.

Lemma 2. If \preceq is change-based then \preceq is reversible if the following holds for its associated relation \sqsubseteq :

$$\text{for any change sets } \gamma, \gamma', \gamma \sqsubseteq \gamma' \text{ if and only if } \text{rev}(\gamma) \sqsubseteq \text{rev}(\gamma')$$

where $\text{rev}(\gamma) = \{\langle \neg f, N - t + 2 \rangle \mid \langle f, t \rangle \in \gamma\}$.

Proof. By definition of rev , $\text{rev}(\text{Ch}(\tau)) = \text{Ch}(\text{reverse}(\tau))$. Now, \preceq is change-based, therefore $\tau < \tau' \Leftrightarrow \text{Ch}(\tau) \sqsubseteq \text{Ch}(\tau')$. Assuming that \sqsubseteq verifies the equivalence $\gamma \sqsubseteq \gamma'$ if and only if $\text{rev}(\gamma) \sqsubseteq \text{rev}(\gamma')$ for any change sets γ and γ' , we get that $\text{rev}(\text{Ch}(\tau)) \sqsubseteq \text{rev}(\text{Ch}(\tau'))$, which is equivalent to $\text{Ch}(\text{reverse}(\tau)) \sqsubseteq \text{Ch}(\text{reverse}(\tau'))$, i.e., $\text{reverse}(\tau) < \text{reverse}(\tau')$. \square

Proposition 6 and Lemma 2 allow us to identify the extrapolation operators among those given in Section 2 that satisfy reversibility.

Proposition 7.

- $E_{nct}, E_{ict}, E_{nc}, E_{ncpt}, E_{ncl}, E_{ncv}, E_{icl}, E_{icv}, E_{ncpl}, E_{ncpv}, E_{cseq}, E_{cseqv}, E_{csi}, E_{csiv}, E_{basic}$ satisfy (R);
- E_k satisfies (R) if and only if $\forall l \in LIT$ we have $k(l) = k(\neg l)$;

Proof. The function rev from change sets to change sets defined by $rev(\langle f, t \rangle) = \langle \neg f, N - t + 2 \rangle$ is bijective, therefore the condition (C) $\gamma \sqsubseteq \gamma'$ if and only if $rev(\gamma) \sqsubseteq rev(\gamma')$ is obviously satisfied for $\preceq_{csi}, \preceq_{csiv}$ and \preceq_{nc} . For every literal l , rev is a one-to-one correspondence between $\gamma(l)$ and $rev(\gamma)(\neg l)$, therefore (C) is satisfied by \preceq_{ncpl} and \preceq_{icl} . For every variable v , rev is a one-to-one correspondence between $\gamma(v)$ and $rev(\gamma)(v)$, therefore (C) is satisfied by \preceq_{ncpv} and \preceq_{icv} . For every time point t , rev is a one-to-one correspondence between $\gamma(t)$ and $rev(\gamma)(N - t + 2)$, therefore (C) is satisfied by \preceq_{ncpt} and \preceq_{ict} . If there is an injective, literal-preserving, time-monotonic function θ from γ to γ' then define θ' from $rev(\gamma)$ to $rev(\gamma')$ as follows: if $\theta(\langle f, t \rangle) = \langle f, t' \rangle$ then $\theta'(\langle \neg f, N - t + 2 \rangle) = \langle \neg f, N - t' + 2 \rangle$. Then θ' is an injective, literal-preserving, time-monotonic function from $rev(\gamma)$ to $rev(\gamma')$, therefore (C) is satisfied by \preceq_{cseq} . The case of \preceq_{cseqv} is similar. Finally, again because for every literal l , rev is a one-to-one correspondence between $\gamma(l)$ and $rev(\gamma)(\neg l)$, we have $k(rev(\gamma)) = \sum_{\langle l, t \rangle \in \gamma} k(\neg l)$, therefore, if the penalty function is such that $k(\neg l) = k(l)$ for any literal l , then $k(rev(\gamma)) = k(\gamma)$, and (C) is satisfied by \preceq_k . \square

If $N > 2$ then E_{chr} does not satisfy (R). Indeed, let $N = 3$, $\mathcal{V} = \{p\}$ and take $\Phi = p_{(1)} \wedge \neg p_{(3)}$, then we have $E_{chr}(\Phi) = p_{(1)} \wedge p_{(2)} \wedge \neg p_{(3)}$ and $E_{chr}(Reverse(\Phi)) = E_{chr}(\neg p_{(1)} \wedge p_{(3)}) = \neg p_{(1)} \wedge \neg p_{(2)} \wedge p_{(3)}$ which differs from $Reverse(E_{chr}(\Phi)) = \neg p_{(1)} \wedge p_{(2)} \wedge p_{(3)}$.

5.2. Markovianity

The meaning of Markovianity is exactly the same as dynamical systems: the way of extrapolating beliefs after time t given that we know the state of the world at time t does not depend on the history of the system before t . Markovianity is also useful from a computational point of view: as soon as we know the state of the world at k different time points, we can decompose the extrapolation process into $k + 1$ smaller independent subprocesses. Transposed in our framework, the property says that as soon as there is a *complete* observation at a given time point k , an extrapolation problem can be decomposed into two independent subproblems: the first one up to k and the second one from k on. For this we need to work with families of extrapolation operators, parameterized by N , rather than for a fixed N , and we slightly abuse notations and use \preceq for $\{\preceq_N, N \in \mathbb{N}\}$, where \preceq_N is a preference relation on $TRAJ_{(N)}$. All “standard” preference relations, including all those that we gave in Section 3, are naturally defined for any N . Similarly, the length of scenarios can now vary.

Notation 1. Let $\Sigma = \langle \Sigma(1), \dots, \Sigma(N) \rangle$ be a scenario of length N and $k \leq N$. Define the scenarios $\Sigma^{\rightarrow k}$ and $\Sigma^{k \rightarrow}$, of respective lengths k and $N - k + 1$, by $\Sigma^{\rightarrow k} = \langle \Sigma(1), \dots, \Sigma(k) \rangle$ and $\Sigma^{k \rightarrow} = \langle \Sigma(k), \dots, \Sigma(N) \rangle$.

Let $TRAJ_{1 \rightarrow N} = TRAJ_{(1)} \cup \dots \cup TRAJ_{(N)}$ be the set of all trajectories of length $\leq N$. If $\tau \in TRAJ_{(N)}$ and $1 \leq k \leq N$ then we denote by $\tau^{\rightarrow k}$ and $\tau^{k \rightarrow}$ the trajectories of respective lengths k and $N - k + 1$ s.t. $\tau^{\rightarrow k} = \langle \tau(1), \dots, \tau(k) \rangle$ and $\tau^{k \rightarrow} = \langle \tau(k), \dots, \tau(N) \rangle$.

A scenario Σ is *complete at k* if $\Sigma(k)$ is complete, i.e., has only one model.

Definition 11 (Markovianity). A scenario-scenario extrapolation operator Ex satisfies Markovianity (M) if and only if for any scenario Σ such that $\Sigma(k)$ is complete,

$$Ex(\Sigma)(t) = \begin{cases} Ex(\Sigma^{\rightarrow k})(t) & \text{if } t \leq k, \\ Ex(\Sigma^{k \rightarrow})(t - k + 1) & \text{if } t \geq k. \end{cases}$$

Note that $Ex(\Sigma)(k)$ is well-defined for $t = k$, because $\Sigma(k)$ is complete, which implies that $Ex(\Sigma)(k) = Ex(\Sigma^{\rightarrow k})(k) = Ex(\Sigma^{k \rightarrow})(1) = \Sigma(k)$.

Example 6. Let $\Sigma = \langle \neg a \vee \neg b, a, a \wedge b, \top, \neg a \rangle$ and $\mathcal{V} = \{a, b\}$. $\Sigma(3)$ is complete. We have $\Sigma^{\rightarrow 3} = \langle \neg a \vee \neg b, a, a \wedge b \rangle$ and $\Sigma^{3 \rightarrow} = \langle a \wedge b, \top, \neg a \rangle$. Assume that $Ex(\Sigma^{\rightarrow 3}) = \langle a \oplus b, a \vee b, a \wedge b \rangle$ and $Ex(\Sigma^{3 \rightarrow}) = \langle a \wedge b, b, \neg a \wedge b \rangle$ (which is the case, for instance, for Ex_{csi}), see Fig. 6. If Ex satisfies (M) then we must have $Ex(\Sigma) = \langle a \oplus b, a \vee b, a \wedge b, b, \neg a \wedge b \rangle$.

We now give a sufficient condition for Markovianity.

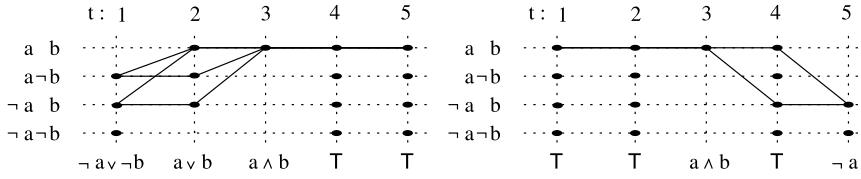


Fig. 6. \preceq_{csi} -preferred trajectories for $\Sigma^{\rightarrow 3}$ and $\Sigma^{3 \rightarrow}$.

Definition 12. A preference relation \preceq on $TRAJ_{(N)}$ is decomposable if there exists a partially ordered set $\langle L, \leq_L \rangle$, a function $\mu : TRAJ_{1 \rightarrow N}$ to L and a function \otimes from $L \times L$ to L , such that:

- for any $\tau \in TRAJ_{(N)}$ and any $k \in \llbracket 1, N \rrbracket$, $\mu(\tau) = \mu(\tau^{\rightarrow k}) \otimes \mu(\tau^{k \rightarrow})$;
- \otimes is strictly increasing (with respect to \leq_L) in both arguments;
- for any $\tau, \tau' \in TRAJ_{(N)}$, $\tau \preceq \tau'$ holds if and only if $\mu(\tau) \leq_L \mu(\tau')$.

Proposition 8. If \preceq is decomposable then Ex satisfies Markovianity.

Proof. Assume \preceq is decomposable and let $\langle L, \leq_L \rangle$, μ and \otimes as above. Let Σ such that $\Sigma(k)$ is complete.

(1) Let $\tau \models Ex(\Sigma)$. Let us show that $\tau^{\rightarrow k} \models Ex(\Sigma^{\rightarrow k})$. Suppose not. Then, using Proposition 1, there exists $\tau_0 \in TRAJ_{(k)}$ such that $\tau_0 \models \Sigma^{\rightarrow k}$ and $\tau_0 < \tau^{\rightarrow k}$; the latter is equivalent to $\mu(\tau_0) <_L \mu(\tau^{\rightarrow k})$. Let τ' be defined by $\tau'(t) = \tau_0(t)$ if $t \leq k$ and $\tau'(t) = \tau(t)$ if $t > k$. Note that $\tau'(k) = \tau_0(k) = \tau(k)$ (since $\tau_0 \models \Sigma^{\rightarrow k}$ and $\tau \models \Sigma$ and $\Sigma(k)$ is complete). Because $\mu(\tau_0) <_L \mu(\tau^{\rightarrow k})$ and \otimes is strictly increasing, $\mu(\tau') = \mu(\tau_0) \otimes \mu(\tau'^{k \rightarrow}) < \mu(\tau^{\rightarrow k}) \otimes \mu(\tau'^{k \rightarrow}) = \mu(\tau^{\rightarrow k}) \otimes \mu(\tau^{k \rightarrow}) = \mu(\tau)$, therefore, $\tau' < \tau$. Lastly, because $\tau_0 \models \Sigma^{\rightarrow k}$ and $\tau'^{k \rightarrow} = \tau^{k \rightarrow} \models \Sigma^{k \rightarrow}$, we have $\tau' \models \Sigma$, which contradicts $\tau \models Ex(\Sigma)$. Therefore, $\tau^{\rightarrow k} \models Ex(\Sigma^{\rightarrow k})$. A similar proof leads to $\tau^{k \rightarrow} \models Ex(\Sigma^{k \rightarrow})$.

(2) Conversely, let $\tau_0 \in TRAJ_{(k)}$ such that $\tau_0 \models Ex(\Sigma^{\rightarrow k})$ and let $\tau_1 \in TRAJ_{(N-k+1)}$ such that $\tau_1 \models Ex(\Sigma^{k \rightarrow})$. Let τ be defined by $\tau(t) = \tau_0(t)$ if $t \leq k$ and $\tau(t) = \tau_1(t - k + 1)$ if $t \geq k$. (Because $\Sigma(k)$ is complete, the definition is consistent at k .) Assume now that $\tau \not\models Ex(\Sigma)$. Then there exists a $\tau' \models \Sigma$ such that $\tau' < \tau$, which entails $\mu(\tau') <_L \mu(\tau)$. We must have either $\mu(\tau'^{\rightarrow k}) <_L \mu(\tau^{\rightarrow k})$ or $\mu(\tau'^{k \rightarrow}) <_L \mu(\tau^{k \rightarrow})$, otherwise by monotonicity of \otimes we would have $\mu(\tau') \geq_L \mu(\tau)$. Assume without loss of generality that $\mu(\tau'^{\rightarrow k}) <_L \mu(\tau^{\rightarrow k})$. This means that $\tau'^{\rightarrow k} < \tau^{\rightarrow k} = \tau_0$. Moreover, $\tau' \models \Sigma$ implies $\tau'^{\rightarrow k} \models \Sigma^{\rightarrow k}$, which together with $\tau'^{\rightarrow k} < \tau_0$ contradicts $\tau_0 \models Ex(\Sigma^{\rightarrow k})$. From (1) and (2), we have that $\tau \models Ex(\Sigma)$ if and only if $\tau^{\rightarrow k} \models Ex(\Sigma^{\rightarrow k})$ and $\tau^{k \rightarrow} \models Ex(\Sigma^{k \rightarrow})$. Let Σ be a scenario such that $\Sigma(k)$ is complete. For every $t \leq k$, we have $Mod(Ex(\Sigma)(t)) = \{\tau(t) \mid \tau \models Ex(\Sigma)\} = \{\tau^{\rightarrow k}(t) \mid \tau^{\rightarrow k} \models Ex(\Sigma^{\rightarrow k})\} = Mod(Ex(\Sigma^{\rightarrow k})(t))$, therefore, $Ex(\Sigma)(t) = Ex(\Sigma^{\rightarrow k})(t)$. Similarly, for $t > k$, $Ex(\Sigma)(t) = Ex(\Sigma^{k \rightarrow})(t - k + 1)$. \square

Proposition 9. The following scenario–scenario extrapolation operators satisfy (M): Ex_{nct} , Ex_{ict} , Ex_{nc} , Ex_{ncpl} , Ex_{cseq} , Ex_{csi} , Ex_{chr} , Ex_k and $Ex_{\mathcal{E}}$.

Proof. In all cases, this is a consequence of Proposition 8, with the following choices for L , μ and \otimes :

- Ex_{nc} (resp. Ex_{nct}): $L = \mathbb{N}$; $\leq_L = \leq$; $\mu(\tau) = |Ch(\tau)|$ (resp. $|ChTime(\tau)|$); and $\otimes = +$. Things are similar for Ex_k with $\mu(\tau)$ being the penalty induced by all changes in τ , and for $Ex_{\mathcal{E}}$ with $\mu(\tau) = \kappa_{\mathcal{E}}(\tau)$.
- Ex_{ncpl} : $L = \mathbb{N}^{LIT}$; for all $x, y \in L$, $x \leq_L y$ if and only if for every $l \in LIT$, $x(l) \leq y(l)$ and $(x \otimes y)(l) = x(l) + y(l)$; $\mu(\tau)(l) = |Ch(\tau)(l)|$. Things are similar for Ex_{ncpv} .
- Ex_{csi} : $L = \langle 2^{LIT \times \{1, \dots, N\}}, \subseteq \rangle$; $\mu(\tau) = Ch(\tau)$; and $\mu(\tau^{\rightarrow k}) \otimes \mu(\tau^{k \rightarrow}) = Ch(\tau^{\rightarrow k}) \cup \{(l, t + k) \mid (l, t) \in Ch(\tau^{k \rightarrow})\}$. Things are similar for Ex_{cseq} and Ex_{chr} , with respectively $\leq_L = \preceq_{cseq}$ and $\leq_L = \preceq_{chr}$. For Ex_{ict} , $L = \langle 2^{\{1, \dots, N\}}, \subseteq \rangle$; $\mu(\tau) = ChTime(\tau)$; and $\mu(\tau^{\rightarrow k}) \otimes \mu(\tau^{k \rightarrow}) = ChTime(\tau^{\rightarrow k}) \cup \{t + k \mid t \in ChTime(\tau^{k \rightarrow})\}$.
- Ex_{cstv} : $L = \langle 2^{\mathcal{V} \times \{1, \dots, N\}}, \subseteq \rangle$; $\mu(\tau) = Ch_u(\tau) = \{(v, t) \mid (v, t) \text{ or } \langle -v, t \rangle \in Ch(\tau)\}$; and $\mu(\tau^{\rightarrow k}) \otimes \mu(\tau^{k \rightarrow}) = Ch_u(\tau^{\rightarrow k}) \cup \{(v, t + k) \mid (v, t) \in Ch_u(\tau^{k \rightarrow})\}$. \square

The property fails typically for operators that use a global minimization like Ex_{icl} (Ex_{ncpt} , Ex_{ncl} , ...). For this kind of operators, the function \otimes is not always strictly increasing. We give a counterexample for Ex_{icl} . Let $\Sigma = \langle a, \neg a, a \wedge b, \neg a \vee \neg b \rangle$, and $k = 3$. There is only one \preceq_{icl} -preferred trajectory satisfying Σ : $\langle ab, \neg ab, ab, \neg ab \rangle$ – the changing literals in this trajectory are a and $\neg a$, and obviously we cannot find anything better. Therefore, $Ex_{icl}(\Sigma) = \langle a \wedge b, \neg a \wedge b, a \wedge b, \neg a \wedge b \rangle$. However, $Ex_{icl}(\Sigma^{\rightarrow 3}) = \langle a \wedge b, a \oplus b \rangle$, therefore, $Ex_{icl}(\Sigma^{\rightarrow 3})(2) \neq Ex_{icl}(\Sigma)(4)$, which shows that Ex_{icl} does not satisfy (M).

5.3. Independence from empty observations

An intuitive property consists in the extrapolation process being insensitive to empty observations, i.e., adding an empty observation between two observations should not change anything to the way observations are extrapolated – or, in other words, the choice of the time unit has no influence on extrapolated beliefs.

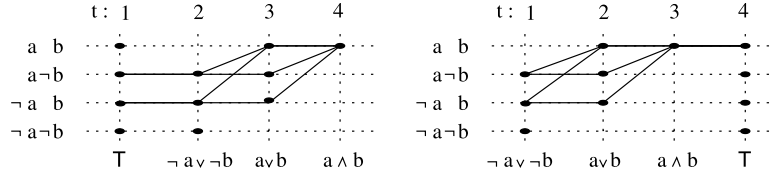


Fig. 7. \preceq_{nc} -preferred trajectories for Σ_1 and Σ_2 .

Definition 13 (Independence from empty observations (IEO)). Ex satisfies (IEO) if for any scenario Σ of length $N - 1$, $\forall p \in \llbracket 1, N \rrbracket$,

$$Ex(\langle \Sigma(1), \dots, \Sigma(p-1), \top, \Sigma(p), \dots, \Sigma(N-1) \rangle) \equiv \langle \varphi^1, \dots, \varphi^{p-1}, \psi, \varphi^p, \dots, \varphi^{N-1} \rangle$$

where $\langle \varphi^1, \dots, \varphi^{N-1} \rangle = Ex(\Sigma)$, and for some propositional formula ψ .

Note that if (IEO) is satisfied, then given a scenario Σ of length $N - 1$, and letting successively $p = N$ and $p = 1$ in the previous identity, we get that $Ex(\langle \Sigma(1), \dots, \Sigma(N-1), \top \rangle) = \langle \varphi^1, \dots, \varphi^{N-1}, \psi^N \rangle$ and $Ex(\langle \top, \Sigma(1), \dots, \Sigma(N-1) \rangle) = \langle \psi^1, \varphi^1, \dots, \varphi^{N-1} \rangle$; this implies that

$$Ex(\langle \Sigma(1), \dots, \Sigma(N), \top \rangle)(t) = Ex(\langle \top, \Sigma(1), \dots, \Sigma(N) \rangle)(t+1)$$

which expresses a kind of *stationarity*: the extrapolation process is invariant to a translation of the time scale.

Example 7. Let $\Sigma_1 = \langle \top, \neg a \vee \neg b, a \vee b, a \wedge b \rangle$ and $\Sigma_2 = \langle \neg a \vee \neg b, a \vee b, a \wedge b, \top \rangle$. We have $Ex_{nc}(\Sigma_1) = \langle a \oplus b, a \oplus b, a \vee b, a \wedge b \rangle$ and $Ex_{nc}(\Sigma_2) = \langle a \oplus b, a \vee b, a \wedge b, a \wedge b \rangle$ see Fig. 7. Hence, $\forall t \in \llbracket 1, 3 \rrbracket, Ex_{nc}(\Sigma_1)(t+1) \equiv Ex_{nc}(\Sigma_2)(t)$. We will see in Proposition 10 that Ex_{nc} indeed satisfies IEO.

We introduce the following notation: for any scenario Σ of length $N - 1$ and any $p \in \llbracket 1, N \rrbracket$, $push(\Sigma, p)$ is the scenario resulting of “pushing” an empty observation \top at time p in Σ :

$$push(\Sigma, p) = \langle \Sigma(1), \dots, \Sigma(p-1), \top, \Sigma(p), \dots, \Sigma(N-1) \rangle, \quad \forall p \in \llbracket 1, N \rrbracket.$$

We now give a useful *sufficient* (but not necessary) condition for an extrapolation operator to satisfy (IEO).

Lemma 3. If \preceq is change-monotonic and satisfies the following two properties

- (1) for any $p \in \llbracket 2, N \rrbracket$ and $\tau \in TRAJ_{(N)}$, $(\tau \rightarrow^{p-1} . \tau(p-1) . \tau^{p+1} \rightarrow) \preceq \tau$;
- (2) for any $\tau_1, \tau_2 \in TRAJ_{(N-1)}$, $\tau_1 \preceq \tau_2$ if and only if $(\tau_1 \rightarrow^{p-1} . \tau_1(p-1) . \tau_1^{p \rightarrow}) \preceq (\tau_2 \rightarrow^{p-1} . \tau_2(p-1) . \tau_2^{p \rightarrow})$

then Ex_{\preceq} satisfies (IEO).

Proof. Let \preceq be a change-monotonic preference relation satisfying (1) and (2), Σ a scenario of length $N - 1$ such that $Ex_{\preceq}(\Sigma) = \langle \varphi^1, \dots, \varphi^{N-1} \rangle$ and $p \in \llbracket 1, N \rrbracket$. Let $\psi = Ex(push(\Sigma, p))(p)$.

We first consider the case $p \geq 2$.

Let $p \geq 2$. We first show that $Ex(push(\Sigma, p)) \models \langle \varphi^1, \dots, \varphi^{p-1}, \psi, \varphi^p, \dots, \varphi^{N-1} \rangle$. Let $\tau \in TRAJ_{(N)}$ such that $\tau \models Ex_{\preceq}(push(\Sigma, p))$, i.e., $\tau \models \langle \Sigma(1), \dots, \Sigma(p-1), \top, \Sigma(p), \dots, \Sigma(N-1) \rangle$ and $\forall \tau' \models push(\Sigma, p)$, $\tau' \not\prec \tau$ (a). Define $\tau_0 \in TRAJ_{(N-1)}$ by $\tau_0 = (\tau \rightarrow^{p-1} . \tau^{p+1} \rightarrow) = \langle \tau(1), \dots, \tau(p-1), \tau(p+1), \dots, \tau(N) \rangle$. We claim that $\tau_0 \models Ex_{\preceq}(\Sigma)$. By definition, $\tau_0 \models \langle \Sigma(1), \dots, \Sigma(p-1), \Sigma(p), \dots, \Sigma(N-1) \rangle = \Sigma$. Assume (b) $\exists \tau'_0 \in TRAJ_{(N-1)}$ such that $\tau'_0 \models \Sigma$ and $\tau'_0 \prec \tau_0$. Define $\tau' \in TRAJ_{(N)}$ by $\tau' = (\tau'_0 \rightarrow^{p-1} . \tau'_0(p-1) . \tau'_0^{p \rightarrow})$; because $\tau'_0 \models \Sigma$, we have $\tau' \models push(\Sigma, p)$. From (2) and (b), we get $\tau' \prec (\tau_0 \rightarrow^{p-1} . \tau_0(p-1) . \tau_0^{p \rightarrow})$. By definition of τ_0 , we have $\tau_0 \rightarrow^{p-1} . \tau_0(p-1) . \tau_0^{p \rightarrow} = \tau \rightarrow^{p-1} . \tau(p-1) . \tau^{p+1} \rightarrow$. Using (1), we get $\tau' \prec (\tau \rightarrow^{p-1} . \tau(p-1) . \tau^{p+1} \rightarrow) \preceq \tau$, which together with $\tau' \models push(\Sigma, p)$ contradicts (a). Thus, (b) cannot hold, i.e., $\forall \tau'_0 \models \Sigma$, $\tau'_0 \not\prec \tau_0$. Hence, $\tau_0 \models Ex_{\preceq}(\Sigma) = \langle \varphi^1, \dots, \varphi^{N-1} \rangle$. Moreover, $\tau_0 = \langle \tau(1), \dots, \tau(p-1), \tau(p+1), \dots, \tau(N) \rangle$, i.e., $\tau \models \langle \varphi^1, \dots, \varphi^{p-1}, \psi, \varphi^p, \dots, \varphi^{N-1} \rangle$.

Still when $p \geq 2$, we show that $\langle \varphi^1, \dots, \varphi^{p-1}, \psi, \varphi^p, \dots, \varphi^{N-1} \rangle \models Ex(push(\Sigma, p))$: we start from a trajectory $\tau_0 \in TRAJ_{(N-1)}$ satisfying $\langle \varphi^1, \dots, \varphi^{N-1} \rangle$ and we claim that $\tau = (\tau_0 \rightarrow^{p-1} . \tau_0(p-1) . \tau_0^{p \rightarrow})$ satisfies $Ex_{\preceq}(push(\Sigma, p))$. Because $\tau_0 \models \langle \varphi^1, \dots, \varphi^{N-1} \rangle$, we have $\tau \models push(\Sigma, p)$. Assume there exists $\tau' \in TRAJ_{(N)}$ such that $\tau' \models push(\Sigma, p)$ and $\tau' \prec \tau$; then we define $\tau'_0 \in TRAJ_{(N-1)}$ by $\tau'_0 = (\tau' \rightarrow^{p-1} . \tau'^{p+1} \rightarrow)$. We have $\tau'_0 \models \Sigma$. Using (1), we get $\tau' \rightarrow^{p-1} . \tau'(p-1) . \tau'^{p+1} \rightarrow \preceq \tau'$. Now, $(\tau' \rightarrow^{p-1} . \tau'^{p+1} \rightarrow) = \tau'_0 \rightarrow^{p-1} . \tau'_0(p-1) . \tau'_0^{p \rightarrow}$. Using (2), we get $\tau'_0 \rightarrow^{p-1} . \tau'_0(p-1) . \tau'_0^{p \rightarrow} \prec \tau_0 \rightarrow^{p-1} . \tau_0(p-1) . \tau_0^{p \rightarrow}$, i.e., $\tau'_0 \prec \tau_0$, which contradicts $\tau_0 \models \langle \varphi^1, \dots, \varphi^{N-1} \rangle = Ex_{\preceq}(\Sigma)$, hence $\tau \models Ex_{\preceq}(push(\Sigma, p))$.

Finally, we consider the case $p = 1$ separately. From the fact that \preceq is change monotonic, two conditions similar to (1) and (2) hold, namely (1') $\forall \tau \in \text{TRAJ}_{(N)}, (\tau(2), \tau^{2 \rightarrow}) \preceq \tau$, and (2') $\forall \tau_1, \tau_2 \in \text{TRAJ}_{(N-1)}, \tau_1 \preceq \tau_2 \Leftrightarrow (\tau_1(1), \tau_1^{1 \rightarrow}) \preceq (\tau_2(1), \tau_2^{1 \rightarrow})$. (1') holds because $\text{Ch}(\tau(2), \tau^{2 \rightarrow}) \subseteq \text{Ch}(\tau)$ and \preceq is change-monotonic. (2') holds because $\text{Ch}(\tau_1(1), \tau_1^{1 \rightarrow}) = \text{Ch}(\tau_1)$ and also $\text{Ch}(\tau_2(1), \tau_2^{1 \rightarrow}) = \text{Ch}(\tau_2)$ and \preceq is change-based. The rest of the proof is similar to the case $p \geq 2$, using (1') and (2') instead of (1) and (2). \square

Proposition 10. *The extrapolation operators based on $\preceq_{nct}, \preceq_{nc}, \preceq_{ncl}, \preceq_{icl}, \preceq_{ncpl}, \preceq_{chr}$ and \preceq_k satisfy (IEO).*

Proof. We first check that condition (1) of Lemma 3 holds for each of these preference relations. Because all of them are changed-based and the two trajectories in (1) coincide except at time p , this amounts to check that $\langle \tau(p-1), \tau(p-1), \tau(p+1) \rangle \preceq \langle \tau(p-1), \tau(p), \tau(p+1) \rangle$. This is easy to check for $\preceq_{nct}, \preceq_{nc}, \preceq_{ncl}, \preceq_{icl}, \preceq_{ncpl}, \preceq_{chr}$ and \preceq_k . Condition (2) holds because the change set of the trajectory in the right side of (2) is the same as the change set of the trajectory in the left side, except that after $p-1$, the changes are shifted by one time unit. \square

A specific proof has to be done for $Ex_{ict}, Ex_{ncpt}, Ex_{cseq}$ and Ex_{csi} , because they do not satisfy condition (1) of Lemma 3. For sake of brevity, we only give the proof for Ex_{csi} (the proofs for the other operators would be similar).

Lemma 4. *Let Σ be a scenario of length $N-1$ and $\tau \in \text{TRAJ}_{(N)}$.*

- for any $p \in \llbracket 2, N \rrbracket$, if $\tau \models Ex_{csi}(\text{push}(\Sigma, p))$ then $(\tau^{\rightarrow p-1}, \tau(p-1), \tau^{p+1 \rightarrow}) \models Ex_{csi}(\text{push}(\Sigma, p))$;
- if $\tau \models Ex_{csi}(\text{push}(\Sigma, 1))$ then $(\tau(2), \tau^{2 \rightarrow}) \models Ex_{csi}(\text{push}(\Sigma, 1))$.

Proof. The second point is straightforward from the fact that $\text{Ch}(\tau(2), \tau^{2 \rightarrow}) \subseteq \text{Ch}(\tau)$.

As for the first point, let $p \in \llbracket 2, N \rrbracket$, and let $\tau \models Ex_{csi}(\text{push}(\Sigma, p))$. Let $\tau_p = (\tau^{\rightarrow p-1}, \tau(p-1), \tau^{p+1 \rightarrow})$, and assume there is a $\tau' \in \text{TRAJ}_{(N)}$ such that $\tau' \models \text{push}(\Sigma, p)$ and $\tau' \prec_{csi} \tau_p$. $\tau' \prec_{csi} \tau_p$ is equivalent to $\text{Ch}(\tau') \subset \text{Ch}(\tau_p)$, which implies that $\tau'(p) = \tau'(p-1)$. Let us define the trajectory τ'' based on τ' and τ as follows: $\forall t \neq p, \tau''(p) = \tau'(p)$ and for any literal $l \in \text{LIT}$, if $\langle l, p+1 \rangle \in \text{Ch}(\tau')$, we set $\tau''(p)(l) = \tau(p)(l)$ else we set $\tau''(p)(l) = \tau''(p-1)(l)$.

It is easy to check that $\tau'' \models \text{push}(\Sigma, p)$. Moreover, by construction, for any time point different from p and $p+1$, $\text{Ch}(\tau'', t) \subseteq \text{Ch}(\tau, t)$. Furthermore, since we assumed that $\tau' \prec_{csi} \tau_p$, any literal l that changes between p and $p+1$ in τ' also changes between p and $p+1$ in τ_p . Since $\tau_p(p-1) = \tau(p-1)$ and $\tau_p(p) = \tau(p-1)$ and $\tau_p(p+1) = \tau(p+1)$, we get that this literal changes between time points $p-1$ and $p+1$ in τ , i.e., $\tau(p-1) \models \neg l$ and $\tau(p+1) \models l$. Hence, either $\tau(p) \models l$, and in this case, by construction of τ'' we get $\langle l, p \rangle \in \text{Ch}(\tau'')$ (and $\langle l, p+1 \rangle \notin \text{Ch}(\tau'')$) and $\langle l, p \rangle \in \text{Ch}(\tau)$; or $\tau(p) \models \neg l$, and in that case, by construction of τ'' we get $\langle l, p+1 \rangle \in \text{Ch}(\tau'')$ and $\langle l, p+1 \rangle \in \text{Ch}(\tau)$. This is true for any literal changing in τ' . The definition of τ'' implies that any change $\langle l, t \rangle$ belonging to $\text{Ch}(\tau'')$ also belongs to $\text{Ch}(\tau')$, except for $t = p+1$, for which the change may either occur at time point p or at $p+1$ accordingly to the changes in τ . Hence any change in τ'' belongs to τ . Now, the strict inclusion comes from the fact that $\text{Ch}(\tau') \subset \text{Ch}(\tau_p)$, since we do not add any change to τ' by defining τ'' but we may only change the time point when it occurs, which means that $\text{Ch}(\tau'') \subset \text{Ch}(\tau)$. Hence, the assumption that there exists a trajectory τ' satisfying $\text{push}(\Sigma, p)$ such that $\tau' \prec_{csi} \tau_p$ cannot hold. Therefore, $\tau_p \models Ex_{csi}(\text{push}(\Sigma, p))$. \square

Proposition 11. *Ex_{csi} satisfies (IEO).*

Proof. For $p \geq 2$, the proof is similar to the proof of Lemma 3. As for the case $p = 1$, the proof is very similar to the proof for the case $p \geq 2$ but based on (1') and (2') given in the proof of Lemma 3, which both hold because \preceq_{csi} is change-monotonic.

Let τ be a \preceq_{csi} -preferred trajectory satisfying $\text{push}(\Sigma, p)$. Define $\tau_0 \in \text{TRAJ}_{(N-1)}$ by $\tau_0 = (\tau^{\rightarrow p-1}, \tau^{p+1 \rightarrow})$; then we have $\tau_0 \models \Sigma$. Assume now that there exists a trajectory τ'_0 such that $\tau'_0 \prec_{csi} \tau_0$ and $\tau'_0 \models \Sigma$. Since \preceq_{csi} satisfies condition (2) of Lemma 3, we have $\tau'_0 \prec_{csi} \tau_0 \Leftrightarrow (\tau'^{\rightarrow p-1}_0, \tau'_0(p-1), \tau'^{p \rightarrow}_0) \prec_{csi} (\tau^{\rightarrow p-1}, \tau(p-1), \tau^{p \rightarrow})$. However, this contradicts Lemma 4, therefore, we get that $\tau_0 \models Ex(\Sigma)$, i.e., $\tau \models \langle \varphi^1, \dots, \varphi^{p-1}, \psi, \dots, \varphi^{N-1} \rangle$.

For the converse direction, we start from a trajectory $\tau_0 \in \text{TRAJ}_{(N-1)}$ such that $\tau_0 \models Ex(\Sigma)$, and then we define $\tau \in \text{TRAJ}_{(N)}$ by $\tau = (\tau_0^{\rightarrow p-1}, \tau_0(p-1), \tau_0^{p \rightarrow})$; we have $\tau \models \text{push}(\Sigma, p)$. Assume there exists $\tau' \in \text{TRAJ}_{(N)}$ such that $\tau' \models \text{push}(\Sigma, p)$ and $\tau' \prec_{csi} \tau$, then we define $\tau'_0 \in \text{TRAJ}_{(N-1)}$ by $\tau'_0 = (\tau'^{\rightarrow p-1}, \tau'^{p+1 \rightarrow})$. We have $\tau'_0 \models \Sigma$. Moreover, we have $\tau = \tau_0^{\rightarrow p-1}, \tau_0(p-1), \tau_0^{p \rightarrow}$, i.e., $\tau(p-1) = \tau(p) = \tau_0(p-1)$. Now, if $\tau(p-1) = \tau(p)$, the same holds for τ' , i.e., $\tau'(p-1) = \tau'(p) = \tau'_0(p-1)$. Due to the fact that \preceq_{csi} satisfies condition (2) of Lemma 3, this implies that $\tau'_0 \prec \tau_0$ which contradicts $\tau_0 \models Ex(\Sigma)$. Therefore, we get that $\tau \models Ex(\text{push}(\Sigma, p))$. \square

We end this section by giving an example of extrapolation operator that does not satisfy IEO. Define \preceq on $\text{TRAJ}_{(N)}$ by $\tau \preceq \tau'$ if and only if $\text{Ch}(\tau, 2) \subseteq \text{Ch}(\tau', 2)$. Then Ex_{\preceq} does not satisfy IEO: take $\Sigma = \langle p \vee q, \neg p \vee \neg q \rangle$, then $Ex_{\preceq}(\Sigma) = \langle p \oplus q, p \oplus q \rangle$ whereas $Ex_{\preceq}(\text{push}(\Sigma, 1)) = Ex_{\preceq}(\langle \top, p \vee q, \neg p \vee \neg q \rangle) = \langle p \vee q, p \vee q, \neg p \vee \neg q \rangle$.

6. Computational issues

6.1. Complexity

We start by identifying the complexity of belief extrapolation for some of the most representative preference relations defined in Section 3. We assume that the reader is familiar with the complexity classes Δ_2^P , Θ_2^P , Π_2^P of the polynomial hierarchy. (Since this is not crucial for the rest of the paper, we do not give any details. The reader can refer to, e.g., [26].)

For each preference relation \preceq_x , we consider the following two decision problems:

- EXTRA_x: given two temporal formulae Φ, Ψ , decide whether $E_x(\Phi) \models \Psi$;
- EXTRA-SC_x: given two scenarios Σ and Σ' , decide whether $Ex_x(\Sigma) \models \Sigma'$.

Proposition 12.

1. EXTRA_x and EXTRA-SC_x are Π_2^P -complete for every $x \in \{csi, icl, cseq, chr\}$;
2. EXTRA_k and EXTRA-SC_k are in Δ_2^P , and are Δ_2^P -complete for some function k ;
3. EXTRA_{nc} and EXTRA-SC_{nc} are $\Delta_2^P(O(\log n))$ -complete.

Point 2 has to be drawn together with the complexity of inference under the backwards semantics in [23] – see Section 8.5.

Lemma 5. For any preference relation \preceq and any two scenarios Σ and Σ' , we have $Ex(\Sigma) \models \Sigma'$ if and only if $E(\Sigma) \models \Sigma'$.

Proof. From left to right: $E(\Sigma) \models Ex(\Sigma)$ (see point 1 of Observation 4), from which the implication follows. From right to left: if $E(\Sigma) \models \Sigma'$ then $S(E(\Sigma)) \models S(\Sigma')$ (S is monotonic with respect to logical consequence, see Observation 2), that is, $Ex(\Sigma) \models \Sigma'$. \square

Therefore, the identity function is a polynomial reduction from EXTRA-SC_x to EXTRA_x, from which we have:

Corollary 12.1. For any x , EXTRA_x is at least as difficult as EXTRA-SC_x.

Therefore, we structure the proof as follows: for each \preceq_x , we show that EXTRA_x belongs to some complexity class C , and then that EXTRA-SC_x is C -hard; this allows us to conclude that both EXTRA_x and EXTRA-SC_x are C -complete.

Lemma 6. EXTRA_{csi}, EXTRA_{icl}, EXTRA_{cseq} and EXTRA_{chr} are in Π_2^P .

Proof. Consider the following nondeterministic algorithm:

1. guess a trajectory $\tau = \langle m_1, \dots, m_n \rangle$;
2. check that $\tau \models \Phi$;
3. check that there is no τ' such that $\tau' \models \Sigma$ and $\tau' \prec \tau$;
4. check that $\tau \models \neg\Psi$.

This algorithm checks that $E(\Phi) \not\models \Psi$. Steps 2 and 4 are done in polynomial time. Step 3 consists of a NP-oracle, since the problem of checking that there is no τ' such that $\tau' \models \Phi$ and $\tau' \prec \tau$ is in coNP as soon as $\tau' \prec \tau$ can be checked in polynomial time, which is the case for all preference relations considered here. Therefore, the algorithm works in nondeterministic polynomial time with NP oracles, hence the membership to Σ_2^P of the complementary problems of EXTRA_{csi}, EXTRA_{icl}, EXTRA_{cseq} and EXTRA_{chr}. These problems are therefore all in Π_2^P . \square

Lemma 7. EXTRA_k is in Δ_2^P .

Proof. Given Φ and an integer Q , the problem (P) of checking whether there exists a trajectory $\tau \models \Phi$ such that $k(\tau) \leq Q$ is in NP, since $k(\tau)$ is computed in polynomial time. Let $M = \sum_{v \in \mathcal{V}} \max(k(v), k(\neg v))$. Let $K^*(\Phi) = \min\{k(\tau) \mid \tau \models \Phi\}$. Clearly, $0 \leq K^*(\Phi) \leq (N-1) \cdot M$, therefore, $K^*(\Phi)$ can be computed using (P) by dichotomy, and the number of calls to (P) in the dichotomous algorithm bounded by $\lceil \log(N \cdot M) \rceil$. Now, $N \cdot M$ is at most exponentially large with respect to the size of the input (where we assume that numbers are, as usual, represented with the binary notation); therefore, the number of calls to (P) is polynomial. Now, consider the following algorithm:

1. compute K^* ;
2. guess τ ;
3. check that $k(\tau) = K^*$;
4. check that $\tau \models \Phi$;
5. check that $\tau \models \neg\Psi$.

This algorithm checks that $Ex_{cp}^k(\Phi) \not\models \Psi$ and works in polynomial time with a polynomial number of calls to NP oracles; hence the membership of the complementary problem of $EXTRA_k$ to Δ_2^P , which shows membership of $EXTRA_k$ to $\Delta_2^P (= co\Delta_2^P)$. \square

Lemma 8. $EXTRA\text{-}SC_{csi}$ is Π_2^P -hard.

Proof. We exhibit a polynomial reduction from the Π_2^P -complete problem SIMPLE BASE REVISION [25]. The latter is defined as follows: given a finite set $\Delta = \{\varphi_1, \dots, \varphi_p\}$ of propositional formulae, and two propositional formulae α and β , let $MaxCons(\Delta, \alpha)$ be the set of all maximal α -consistent subsets of Δ (that is, $S \in MaxCons(\Delta, \alpha)$ if and only if $S \cup \{\alpha\}$ is consistent and there is no $S' \subseteq \Delta$ such that $S' \supset S$ and $S' \cup \{\alpha\}$ is consistent). Then $\langle \Delta, \alpha, \beta \rangle$ is a positive instance of SIMPLE BASE REVISION if and only if for each $S \in MaxCons(\Delta, \alpha)$ we have $S \cup \{\alpha\} \models \beta$.

Let us now exhibit the polynomial reduction F from SIMPLE BASE REVISION to $EXTRA_{csi}$: any instance $\langle \Delta, \alpha, \beta \rangle$ of SIMPLE BASE REVISION is mapped to the instance of $EXTRA_{csi}$ defined by

- $N = 2$;
- $Var(\Sigma) = Var(\Delta) \cup Var(\alpha) \cup Var(\beta) \cup \{x_1, \dots, x_p\}$, where $Var(\varphi)$ is the set of variables appearing in the formula φ and x_1, \dots, x_p are new variables (not appearing in Δ, α and β);
- $\Sigma(1) = x_1 \wedge \dots \wedge x_p$;
- $\Sigma(2) = \alpha \wedge (\varphi_1 \leftrightarrow x_1) \wedge \dots \wedge (\varphi_p \leftrightarrow x_p)$;
- $\Sigma' = \langle \top, \beta \rangle$.

We now prove that $Ex_{csi}(\Sigma)(2) \models \beta$ if and only if $\langle \Delta, \alpha, \beta \rangle$ is a positive instance of SIMPLE BASE REVISION.

Let $X = \{x_1, \dots, x_p\}$ and $Y = Var(\Delta) \cup Var(\alpha) \cup Var(\beta)$. We say that a trajectory τ is Y -static if for any variable $y \in Var(\Delta) \cup Var(\alpha) \cup Var(\beta) \setminus \{y\}$, we have $\tau(1)(y) = \tau(2)(y)$.

Let $\tau \in Min(\preceq_{csi}, Traj(\Sigma))$. We first show that τ is Y -static. Suppose it is not the case, i.e., there is an $y \in Y$ such that $\tau(1)(y)$ and $\tau(2)(y)$ differ. We then build the following trajectory τ' from τ as follows:

- $\tau'(2) = \tau(2)$;
- for each $x \in X \cup Y$, $\tau'(1)(x) = \tau(1)(x)$;
- $\tau'(1)(y) = \tau(2)(y)$.

We have $\tau' \models \Sigma'$: indeed, $\tau'(2) = \tau(2)$ and $\tau \models \Sigma$ imply $\tau'(2) \models \Sigma'$, and since y does not appear in $\Sigma(1)$, the truth value of y in $\tau(1)$ has no influence on the satisfaction of $\Sigma(1)$, hence, $\tau(1) \models \Sigma(1)$ implies $\tau'(1) \models \Sigma(1)$. Now, $Ch(\tau')$ is strictly contained in $Ch(\tau)$, since it contains the changes of τ except the change regarding y . This contradicts $\tau \in Min(\preceq_{csi}, Traj(\Sigma))$. Therefore, any $\tau \in Min(\preceq_{csi}, Traj(\Sigma))$ is such that $\tau(1)(y) = \tau(2)(y)$ for any $y \in Y$.

Now, there is a one-to-one correspondence between Y -static trajectories in $Traj(\Sigma)$ and α -consistent subsets of Δ . Let $\tau \in Traj(\Sigma)$ and let $SB(\tau) = \{\varphi_i \mid \tau(2) \models \varphi_i\} = \{\varphi_i \mid \langle x_i, 2 \rangle \notin Ch(\tau)\}$. If τ is Y -static then $SB(\tau)$ is α -consistent, because $\tau(1) \models \bigwedge_{\varphi_i \in SB(\tau)} (x_i \wedge \varphi_i) \wedge \alpha$, therefore $\tau(1) \models SB(\tau)$. Conversely, if B is a α -consistent subset of Δ then let $m \models B \wedge \alpha$ and build the trajectory τ_B as follows: $\tau_B(1)(v) = m(v), \forall v \in Y$ and $\tau_B(1)(v) = true \forall v \in \{x_1, \dots, x_p\}$ and $\tau_B(2)(v) = m(v), \forall v \in Y$ and $\tau_B(2)(x_i) = true$ iff $m \models \varphi_i$. $\tau_B(1)$ and $\tau_B(2)$ coincide on all variables of Y , therefore τ_B is Y -static. Moreover, we have $\tau_B(1) \models \Sigma(1)$, and $\tau_B(2) \models \Sigma(2)$ (the latter holds because $\tau_B(2) \models \alpha$ as a consequence of $m \models \alpha$), and because $\tau_B(2) \models \varphi_i$ if and only if $\tau_B(2) \models x_i$ (by construction of τ). Therefore, $\tau_B \models \Sigma$.

Now, $\tau \in Min(\preceq_{csi}, Traj(\Sigma))$ implies that $\{x_i \mid \tau(1) \models \neg x_i\}$ is minimal, that is, there is no $\tau' \in Min(\preceq_{csi}, Traj(\Sigma))$ such that $\{x_i \mid \tau'(1) \models \neg x_i\} \subset \{x_i \mid \tau(1) \models \neg x_i\}$. We now have the following chain of equivalences:

$$\begin{aligned} \{x_i \mid \tau(2) \models \neg x_i\} \text{ is minimal} \\ \Leftrightarrow \{x_i \mid \tau(2) \models \neg \varphi_i\} \text{ is minimal} \\ \Leftrightarrow \{x_i \mid \tau(2) \models \varphi_i\} \text{ is maximal} \\ \Leftrightarrow \tau(2) \text{ satisfies a maximal } \alpha\text{-consistent subset of } \Delta. \end{aligned}$$

Lastly, we have $\tau(2) \models \alpha$. Therefore, we have $\Delta \star \alpha \models \beta$ if and only if $Ex_{csi}(\Sigma)(2) \models \beta$. \square

Lemma 9. $EXTRA\text{-}SC_k$ is Δ_2^P -hard for some penalty function k .

Proof. By reduction from LINEAR BASE REVISION [25]. The proof uses a reduction similar to the reduction used in the proof of Lemma 8: the scenarios Σ and Σ' are defined exactly in the same way, only the criterion for maximality changes. Consider the penalty-based extrapolation operator defined by the weights $k(x_i) = 2^{n-i+1}$ for every $i = 1, \dots, n$. Let Δ , α and β be defined as in the proof of Lemma 8. The linear base revision of Δ by α is defined by $\Delta \star_L \alpha = Cn(\Delta' \cup \{\alpha\})$, where Δ' is defined inductively as follows:

- $\Delta'_0 = \emptyset$;
- for all $i = 1, \dots, p$, $\Delta'_i = \Delta'_{i-1} \cup \{\varphi_i\}$ if $\Delta'_{i-1} \cup \{\varphi_i\} \cup \{\alpha\}$ is satisfiable, and $\Delta'_i = \Delta'_{i-1}$ otherwise;
- $\Delta' = \Delta'_n$.

Let Σ and Σ' be defined as in the proof of Lemma 8, as well as the correspondence between Y -static trajectories and α -consistent subsets of Δ . Using the very same line of argumentation as in Lemma 8, τ is a Y -trajectory minimizing $k(\tau)$ if and only if $\tau(2) \models \Delta'$, from which we get $\Delta \star_L \alpha \models \beta$ if and only if $Ex_k(\Sigma)(2) \models \beta$. \square

Lemma 10. $EXTRA_{nc}$ is in $\Delta_2^P(O(\log n))$.

Proof. The proof is essentially similar to the membership proof of $EXTRA_k$ to Δ_2^P , except that since trajectory costs are the number of changes, the number of possible trajectory costs is polynomial (namely, in $o(N \cdot |\mathcal{V}|)$) and the computation of $K^*(\Sigma)$ therefore only requires a logarithmic number (in the size of the input) of calls to (P). \square

Lemma 11. $EXTRA-SC_{nc}$ is $\Delta_2^P(O(\log n))$ -hard.

Proof. By reduction from CARDINALITY-MAXIMIZING BASE REVISION [25]. Let Δ , α and β be defined as in the proof of Lemma 8. The cardinality-maximizing base revision of Δ by α is defined by $\Delta \star_C \alpha = \bigcap \{Cn(\Delta' \cup \{\alpha\}), \Delta' \in MaxCard(B, \alpha)\}$, where $MaxCard(\Delta, \alpha)$ is the set of all subsets of B of maximal cardinality that are consistent with α .

Again, the proof uses the same reduction as in the proof of Lemma 8. Let Σ and Σ' be defined as in the proof of Lemma 8, as well as the correspondence between Y -static trajectories and α -consistent subsets of Δ . We easily check that the number of changes in τ is equal to $p - |B_\tau|$, therefore, \preceq_{nc} -preferred trajectories of τ correspond to α -consistent subsets of B of maximum cardinality, from which we get $\Delta \star_C \alpha \models \beta$ if and only if $Ex_{nc}(\Sigma)(2) \models \beta$. \square

Lemma 12. $EXTRA-SC_{icl}$, $EXTRA-SC_{cseq}$ and $EXTRA-SC_{chr}$ are Π_2^P -hard.

Proof. This is a straightforward corollary of Lemma 8, once noticed that when $N = 2$, the preference relations \preceq_{icl} , \preceq_{cseq} and \preceq_{chr} all coincide with \preceq_{csi} . \square

Proposition 12 follows from the previous lemmas and Corollary 12.1.

6.2. Computing extrapolated beliefs

In this section, we focus on the practical computation of extrapolated beliefs. A bad news is that techniques for computing extrapolation are generally sensitive to the choice of the preference relation, which means that, to a large extent, the study has to be done independently for each preference relation. Considering all preference relations introduced in Section 2 would be far too long. Rather, we focus on the “prototypical” preference relation \preceq_{csi} , and on scenario–scenario extrapolation. At the end of the section we briefly say to what extent our results can be generalized to other preference relations, or to formula–formula extrapolation.

Accordingly, for the rest of this section, we fix the preference relation to \preceq_{csi} and we show that the computation of the preferred trajectories of a given scenario Σ w.r.t. \preceq_{csi} (corresponding to the set of what we call *minimal explanations for this scenario*) can be done in four steps:

1. for each variable, compute its *relevant time points* for Σ ;
2. generate the *persistence formulae* (of the form $v_{(t)} \leftrightarrow v_{(t')}$) associated to each variable w.r.t. its relevant time points (computed at step 1);
3. select the *maximal subsets of persistence formulae consistent with Σ* ;
4. compute the *minimal change sets* corresponding to the maximal persistence sets (obtained at step 3).

Definition 14.

- A change set $\gamma = \{\langle l_1, t_1 \rangle, \dots, \langle l_k, t_k \rangle, l_i \in LIT, t_i \in \llbracket 2, N \rrbracket\}$ explains a scenario Σ if there exists a $\tau \in Traj(\Sigma)$ such that $Ch(\tau) = \gamma$.
- A *minimal explanation* γ for Σ is the change set $Ch(\tau)$ of a \preceq_{csi} -preferred trajectory τ of Σ .

Intuitively, a change set explains Σ if it is consistent with it; this change set is a minimal explanation for Σ if it corresponds to the change set of a preferred trajectory of Σ , and therefore “explains” the scenario Σ .

Given a scenario Σ (of length N), the number of minimal explanations can be very large, even when the number of changes is small:

Example 8. $\Sigma^0 = \langle a \wedge b \wedge c \wedge d, \top, \neg b, d, \neg a, \top, b, \neg c \vee \neg d \rangle$ has 352 minimal explanations (and as many minimal trajectories). Now, these 352 minimal explanations can be expressed compactly: γ is a minimal explanation for Σ^0 if and only if it contains exactly the following changes:

- (i) $\langle \neg a, t \rangle$ for exactly one $t \in \{2..5\}$;
- (ii) $\langle \neg b, t \rangle$ for exactly one $t \in \{2, 3\}$;
- (iii) $\langle b, t \rangle$ for exactly one $t \in \{4..7\}$;
- (iv) either (iv') $\langle \neg c, t \rangle$ for one $t \in \{2..8\}$ or (iv'') $\langle \neg d, t \rangle$ for one $t \in \{5..8\}$.

Thus we get a total of $4 \times 2 \times 4 \times (7 + 4) = 352$ minimal explanations.

To capture this we introduce the notion of *compact change sets*, which allow for representing succinctly a set of change “contiguous” sets.

Definition 15 (*Compact change sets*).

- A compact change set Γ is a set $\{\langle l_i, t_i^-, t_i^+ \rangle \mid i = 1..p\}$ where each l_i is a literal and each (t_i^-, t_i^+) is a pair of time points such that $0 < t_i^- \leq t_i^+ \leq N$.
- A change set γ is covered by a compact change set $\Gamma = \{\langle l_i, t_i^-, t_i^+ \rangle \mid i = 1..p\}$ if $\gamma = \{\langle l'_1, t_1 \rangle, \dots, \langle l'_p, t_p \rangle\}$, and there is a bijection mapping each $\langle l'_i, t_i \rangle \in \gamma$ to a unique $\langle l_i, t_i^-, t_i^+ \rangle \in \Gamma$ such that $l'_i = l_i$ and $t_i^- \leq t_i \leq t_i^+$. $Cov(\Gamma)$ denotes the set of all change sets covered by Γ .

Let Γ be a compact change set and Σ a scenario. We say that

- Γ explains Σ if and only if every $\gamma \in Cov(\Gamma)$ explains Σ ,
- Γ explains minimally Σ if and only if $\forall \gamma \in Cov(\Gamma)$, γ is a minimal explanation for Σ .

Intuitively, a compact change set consists of the disjunction of all the change sets it covers.

Example 8 (*Continued*). $\Gamma = \{\langle \neg a, 2, 5 \rangle, \langle \neg b, 2, 3 \rangle, \langle b, 4, 7 \rangle, \langle \neg c, 2, 8 \rangle\}$ explains $\Sigma^0 = \langle a \wedge b \wedge c \wedge d, \top, \neg b, d, \neg a, \top, b, \neg c \vee \neg d \rangle$. It is a compact change set covered by $4 \times 2 \times 4 \times 7 = 224$ minimal explanations such as, for instance, $\{\langle \neg a, 2 \rangle, \langle \neg b, 2 \rangle, \langle \neg c, 2 \rangle, \langle b, 5 \rangle\}$.

We now define the notion of *maximally covering minimal explanation* (MCME). The set of MCMEs for Σ expresses minimal explanations in a compact, redundancy-free way.

Definition 16. Γ is a *maximally covering minimal explanation* (MCME) for Σ if and only if (i) Γ explains minimally Σ and (ii) there is no Γ' explaining minimally Σ such that $Cov(\Gamma) \subset Cov(\Gamma')$.

Example 8 (*Continued*). Γ is a MCME for Σ^0 . Let $\Gamma' = \{\langle \neg a, 2, 5 \rangle, \langle \neg b, 2, 3 \rangle, \langle \neg c, 4, 6 \rangle, \langle b, 4, 7 \rangle\}$. Γ' minimally explains Σ^0 as well, but it is not maximally covering, because $Cov(\Gamma) \subset Cov(\Gamma')$.

The fact that the set of MCMEs for Σ is redundancy-free is expressed by the following result.

Proposition 13. For each minimal explanation γ for Σ there exists a MCME Γ for Σ covering γ .

Proof. From Definition 16, (a) if Γ explains minimally Σ and is not a MCME for Σ then there exists Γ' explaining minimally Σ such that $Cov(\Gamma) \subset Cov(\Gamma')$.

Let Γ explaining minimally Σ . If Γ is not a MCME for Σ then we apply (a) repeatedly and build a finite sequence $\Gamma_0 = \Gamma, \Gamma_1, \dots, \Gamma_p$ such that for each $i \geq 1$, Γ_i explains minimally Σ and $Cov(\Gamma_{i-1}) \subset Cov(\Gamma_i)$, and we stop until it is not possible to find any Γ_{p+1} such that $Cov(\Gamma_p) \subset Cov(\Gamma_{p+1})$. Since the number of change sets is finite, this process is guaranteed to stop, and the last Γ_p obtained is a MCME. Therefore, (b) if Γ explains minimally Σ then there exists a $\Gamma' \in MCME(\Sigma)$ such that $Cov(\Gamma) \subseteq Cov(\Gamma')$.

Let γ be a minimal explanation for Σ . Applying (b) to $\Gamma = \{\langle l_i, t_i, t_i \rangle \mid \langle l_i, t_i \rangle \in \gamma\}$ enables us to show the existence of a MCME Γ' for Σ covering γ . \square

For the scenario Σ^0 considered in Example 8, there are two MCMES: $\{(\neg a, 2, 5), (\neg b, 2, 3), (b, 4, 7), (\neg c, 2, 8)\}$ and $\{(\neg a, 2, 5), (\neg b, 2, 3), (b, 4, 7), (\neg d, 5, 8)\}$.

When observations are disjunction-free, we only have one maximally compact minimal explanation:

Proposition 14. *If for each t , $\Sigma(t)$ is a conjunction of literals, then there is only one maximally compact minimal explanation for Σ .*

Proof. Let $\Sigma = \langle \Sigma^1, \dots, \Sigma^N \rangle$ where each Σ^i is a consistent conjunction of literals of LIT. For any literal l , let $Int(l)$ be the set of intervals defined by: $\theta = [t^-, t^+] \in Int(l)$ if and only if

- $\Sigma^{t^-} \models \neg l$;
- $\Sigma^{t^+} \models l$;
- for all t such that $t^- < t < t^+$, neither l nor $\neg l$ is a conjunct of Σ^t .

Now, let $\Gamma = \{ \langle l, \theta \rangle \mid l \in LIT, \theta \in Int(l) \}$. Clearly enough, Γ is a MCME for Σ . Since any minimal explanation γ is covered by Γ , there is no other MCME for Σ . \square

Compact minimal explanations can be computed by adding to Σ a maximal set of *persistence formulae* expressing that the value of a variable v persists between two consecutive time points. Now, this would generate as much as $(N - 1) \cdot p$ persistence formulae, where p is the number of propositional variables. Rather than considering *all* persistence formulae, it suffices to consider, for each propositional variable v , formulae expressing that v persists between two consecutive observations that are “relevant” to v . This notion of relevance can be expressed in this simple way (see [21]):

Definition 17.

- A formula φ *depends* on a variable v if and only if every formula equivalent to φ contains an occurrence of v . $DepVar(\varphi)$ denotes the set of all variables on which φ depends.⁹
- For each variable v , define its set of

relevant time points with respect to Σ , denoted by $R_\Sigma(v)$, as the set of all time points $t \in \llbracket 1, N \rrbracket$ such that $v \in DepVar(\Sigma(t))$ and

relevant intervals $RI_\Sigma(v)$ to v with respect to Σ as the set of all intervals $[t, t']$ such that $t, t' \in R_\Sigma(v)$, $t \neq t'$, and $[t + 1, t' - 1] \cap R_\Sigma(v) = \emptyset$.

Thus, t is a relevant time point to v if $\Sigma(t)$ says something about v , and the intuitive meaning of relevant intervals is that it is sound to consider only persistence assumptions regarding a given variable within these intervals only.¹⁰ The persistence set P_Σ is the collection of all such persistence formulae.

Let us illustrate this definition by an example.

Example 9. Let $\mathcal{V} = \{a, b, c, d\}$, $N = 5$ and $\Sigma^1 = \langle \neg a, a \vee c, b \wedge d, \neg a \vee \neg b, \neg c \rangle$.

$R_{\Sigma^1}(a) = \{1, 2, 4\}$, $R_{\Sigma^1}(b) = \{3, 4\}$, $R_{\Sigma^1}(c) = \{2, 5\}$ and $R_{\Sigma^1}(d) = \{3\}$;

$RI_{\Sigma^1}(a) = \{[1, 2], [2, 4]\}$, $RI_{\Sigma^1}(b) = \{[3, 4]\}$, $RI_{\Sigma^1}(c) = \{[2, 5]\}$, $RI_{\Sigma^1}(d) = \emptyset$.

Definition 18. Let Σ be a consistent scenario. We define:

- $\Phi_\Sigma = \bigwedge_{t=1}^N \Sigma(t)$ the temporal formula associated to the scenario Σ ;
- $P_\Sigma = \bigcup \{v_{(t)} \leftrightarrow v_{(t')} \mid v \in \mathcal{V}, [t, t'] \in RI_\Sigma(v)\}$ is the *persistence set* associated with Σ .
- π is a *maximal consistent persistence set* for Σ if and only if it is a maximal subset of P_Σ consistent with Φ_Σ ;
- the *compact change set* $\Gamma(\pi)$ associated with $\pi \subseteq P_\Sigma$ is defined by:

$$\Gamma(\pi) = \left\{ \langle l, [t^-, t^+] \rangle \mid \begin{array}{l} v_{(t^-)} \leftrightarrow v_{(t^+)} \in P_\Sigma \setminus \pi \text{ and} \\ l = \begin{cases} v & \text{if } \pi \cup \{\Phi_\Sigma\} \models v_{(t^+)} \\ \neg v & \text{else} \end{cases} \end{array} \right\}.$$

Example 9 (Continued). Again considering $\Sigma^1 = \langle \neg a, a \vee c, b \wedge d, \neg a \vee \neg b, \neg c \rangle$, these definitions give:

⁹ Equivalently, $v \notin DepVar(\varphi)$ does not depend on v if and only if $\varphi_{v \leftarrow \top}$, $\varphi_{v \leftarrow \perp}$ and φ are logically equivalent [21]. For instance, $DepVar((a \vee \neg a) \wedge \neg b) = \{b\}$.

¹⁰ This method could be further improved by distinguishing between *positive* and *negative* dependence [21].

$$\Phi_{\Sigma^1} = \neg a_{(1)} \wedge (a_{(2)} \vee c_{(2)}) \wedge b_{(3)} \wedge d_{(3)} \wedge (\neg a_{(4)} \vee \neg b_{(4)}) \wedge \neg c_{(5)};$$

$$P_{\Sigma^1} = \{a_{(1)} \leftrightarrow a_{(2)}, a_{(2)} \leftrightarrow a_{(4)}, b_{(3)} \leftrightarrow b_{(4)}, c_{(2)} \leftrightarrow c_{(5)}\}.$$

There are 3 maximal consistent persistence sets for Σ^1 :

$$\pi^1 = \{a_{(2)} \leftrightarrow a_{(4)}, c_{(2)} \leftrightarrow c_{(5)}\};$$

$$\pi^2 = \{b_{(3)} \leftrightarrow b_{(4)}, c_{(2)} \leftrightarrow c_{(5)}\};$$

$$\pi^3 = \{a_{(1)} \leftrightarrow a_{(2)}, a_{(2)} \leftrightarrow a_{(4)}, b_{(3)} \leftrightarrow b_{(4)}\}.$$

$\Gamma(\pi^1) = \Gamma^1 = \{\langle a, 2, 2 \rangle, \langle b, 4, 4 \rangle\}$; $\Gamma(\pi^2) = \Gamma^2 = \{\langle a, 2, 2 \rangle, \langle \neg a, 3, 4 \rangle\}$; $\Gamma(\pi^3) = \Gamma^3 = \{\langle \neg c, 3, 5 \rangle\}$. Their minimal change sets are $\{\langle a, 2 \rangle, \langle b, 4 \rangle\}$ covered by Γ^1 ; $\{\langle a, 2 \rangle, \langle \neg a, 3 \rangle\}$ and $\{\langle a, 2 \rangle, \langle \neg a, 4 \rangle\}$ both covered by Γ^2 ; $\{\langle \neg c, 3 \rangle\}$, $\{\langle \neg c, 4 \rangle\}$ and $\{\langle \neg c, 5 \rangle\}$ covered by Γ^3 .

Example 8 (Continued). Concerning Σ^0 , we obtain the following results: $P_{\Sigma^0} = \{a_{(1)} \leftrightarrow a_{(5)}, b_{(1)} \leftrightarrow b_{(3)}, b_{(3)} \leftrightarrow b_{(7)}, c_{(1)} \leftrightarrow c_{(8)}, d_{(1)} \leftrightarrow d_{(4)}, d_{(4)} \leftrightarrow d_{(8)}\}$. There are two maximal consistent persistence sets for Σ^0 : $\{c_{(1)} \leftrightarrow c_{(8)}, d_{(1)} \leftrightarrow d_{(4)}\}$ and $\{d_{(1)} \leftrightarrow d_{(4)}, d_{(4)} \leftrightarrow d_{(8)}\}$. Their compact change sets are the two MCME given before, namely, $\{\langle \neg a, 2, 5 \rangle, \langle \neg b, 2, 3 \rangle, \langle b, 4, 7 \rangle, \langle \neg c, 2, 8 \rangle\}$ and $\{\langle \neg a, 2, 5 \rangle, \langle \neg b, 2, 3 \rangle, \langle b, 4, 7 \rangle, \langle \neg d, 5, 8 \rangle\}$.

We now have the following result:

Proposition 15. γ is a minimal explanation for Σ if and only if there exists a maximal persistence set π for Σ such that $\gamma \in \text{Cov}(\Gamma(\pi))$.

Lemma 13. If π is a maximal persistence set for Σ then any change set $\gamma \in \text{Cov}(\Gamma(\pi))$ explains minimally Σ .

Proof. Let π be a maximal persistence set for Σ , therefore $\Phi_{\Sigma} \wedge \pi$ is consistent. Let τ_{π} be an interpretation of TRAJ_N such that $\tau_{\pi} \models \Phi_{\Sigma} \wedge \pi$. Obviously, $\tau_{\pi} \models \Phi_{\Sigma}$ is equivalent to $\tau_{\pi} \models \Sigma$. $\tau_{\pi} \models \pi$ entails $\tau_{\pi}(t^{-})(v) = \tau_{\pi}(t^{+})(v)$ for any $v \in \mathcal{V}$ and any $t^{+}, t^{-} \in \llbracket 1, N \rrbracket$ such that $v_{(t^{-})} \leftrightarrow v_{(t^{+})} \in \pi$.

Let γ be a change set covered by $\Gamma(\pi)$. Let us now build τ_{γ} as follows: for each variable $v \in \mathcal{V}$,

1. for all $t \in R_{\Sigma}(v)$, $\tau_{\gamma}(t)(v) = \tau_{\pi}(t)(v)$;
2. for all $\theta = [t^{-}, t^{+}] \in RI_{\Sigma}(v)$:
 - if $\tau_{\pi}(t^{-})(v) = \tau_{\pi}(t^{+})(v)$ then for all $t \in \theta$, we let $\tau_{\gamma}(t)(v) = \tau_{\pi}(t^{-})(v)$ (we impose that τ_{γ} has no change in this interval);
 - if $\tau_{\pi}(t^{-})(v) \neq \tau_{\pi}(t^{+})(v)$, which entails $v_{(t^{-})} \leftrightarrow v_{(t^{+})} \notin \pi$ - if $\tau_{\pi}(v)(t^{+})$ is true (resp. false) then, by definition, $\Gamma(\pi)$ contains $\langle v, t^{-} + 1, t^{+} \rangle$ (resp. $\langle \neg v, t^{-} + 1, t^{+} \rangle$). By definition of Cov , there is one and only one t^{*} in $[t^{-} + 1, t^{+}]$ such that $\langle \neg v, t^{*} \rangle$ (resp. $\langle v, t^{*} \rangle$) is in γ . Then we let

$$\begin{aligned} \tau_{\gamma}(t^{-})(v) &= \dots = \tau_{\gamma}(t^{*} - 1)(v) = \tau_{\pi}(t^{-})(v) \quad \text{and} \\ \tau_{\gamma}(t^{*})(v) &= \dots = \tau_{\gamma}(t^{+})(v) = \tau_{\pi}(t^{+})(v); \end{aligned}$$

3. for all $t < \min(R_{\Sigma}(v))$, we let $\tau_{\gamma}(t)(v) = \tau_{\pi}(\min(R_{\Sigma}(v)))(v)$;
4. for all $t > \max(R_{\Sigma}(v))$, we let $\tau_{\gamma}(t)(v) = \tau_{\pi}(\max(R_{\Sigma}(v)))(v)$.

Because $\tau_{\gamma}(v)$ coincides with $\tau_{\pi}(v)$ for every variable v and any time point $t \in R_{\Sigma}(v)$, we have τ_{γ} gives the same value as τ_{π} to the observations $\varphi_1, \dots, \varphi_N$, therefore $\tau_{\gamma} \models \Sigma$. Moreover, by construction $\text{Ch}(\tau_{\gamma}) = \gamma$.

Let us now assume that τ_{γ} is not a *preferred* trajectory in $\text{Traj}(\Sigma)$ w.r.t. \preceq_{csi} : in this case, there exists a $\tau' \models \Sigma$ such that $\gamma' = \text{Ch}(\tau') \subset \text{Ch}(\tau_{\gamma}) = \gamma$. $\gamma' \subset \gamma$ implies that there is a $v \in \mathcal{V}$ and a $\theta \in RI(v)$ such that γ contains either $\langle v, t \rangle$ or $\langle \neg v, t \rangle$ for $t \in \theta$ and γ' contains none. Let $\theta = [t^{-}, t^{+}]$ and $\pi' = \pi \cup \{v_{(t^{-})} \leftrightarrow v_{(t^{+})}\}$. We have $\tau' \models \Phi_{\Sigma}$ (because $\tau' \models \Sigma$) and $\tau' \models \pi'$ (because γ' explains Σ). Therefore, π' is a persistence set of Σ , which contradicts the assumption that π is a minimal persistence set of Σ . Therefore γ explains minimally Σ . \square

Lemma 14. For any change set γ , if γ is a minimal explanation for Σ then there exists a maximal persistence set π for Σ such that $\gamma \in \text{Cov}(\Gamma(\pi))$.

Proof. Let $\gamma = \{\langle l_1, t_1 \rangle, \dots, \langle l_p, t_p \rangle\}$ be a minimal explanation for Σ . Let us consider the set $\pi_{\gamma} = \{v_{(t^{-})} \leftrightarrow v_{(t^{+})} \in P_{\Sigma} \mid \forall t \in [t^{-} + 1, t^{+}], \langle v, t \rangle$ and $\langle \neg v, t \rangle$ are not in $\gamma\}$. γ being a minimal explanation means that it exists a trajectory $\tau_{\gamma} \in \min(\preceq_{\text{csi}}, \text{Mod}(\Sigma))$ such that $\text{Ch}(\tau_{\gamma}) = \gamma$.

- Let us show now that π_γ is a maximal persistence set for Σ . $\pi_\gamma \cup \{\Phi_\Sigma\}$ is consistent since τ_γ satisfies it. Suppose that π_γ is not maximal, that is, there exists $\pi \supset \pi_\gamma$ such that $\pi \wedge \Phi_\Sigma$ is consistent. Let $\tau \models \pi \wedge \Phi_\Sigma$, then $\tau \models \Sigma$. Now, let change τ into τ' the following way:
 - for each $v \in \mathcal{V}$ and each $t \in R_\Sigma(v)$, $\tau'(t)(v) = \tau(t)(v)$;
 - for each $v \in \mathcal{V}$ and each $[t^-, t^+] \in RI_\Sigma(v)$ such that $v_{(t^-)} \leftrightarrow v_{(t^+)}$ is in π , $\tau'(t)(v) = \tau(t^-)(v) = \tau(t^+)(v)$;
 - for each $v \in \mathcal{V}$ and each $[t^-, t^+] \in RI_\Sigma(v)$ such that $v_{(t^-)} \leftrightarrow v_{(t^+)}$ is not in π , note that $v_{(t^-)} \leftrightarrow v_{(t^+)}$ is not in π_γ either, and therefore there exists a t^* such that $t^- < t^* \leq t^+$ and $\langle v, t^* \rangle$ or $\langle \neg v, t^* \rangle$ is in γ . Then let

$$\tau'(t)(v) = \begin{cases} \tau(t^-)(v) & \text{if } t < t^*, \\ \tau(t^+)(v) & \text{if } t \geq t^*. \end{cases}$$

Since τ' coincides with τ for each $v \in \mathcal{V}$ and each $t \in R_\Sigma(v)$, we have $\tau' \models \Sigma$. Now, $Ch(\tau')$ is the subset of $\gamma = Ch(\tau_\gamma)$ consisting of all pairs $\langle v, t \rangle$ or $\langle \neg v, t \rangle$ of γ such that $v_{(t^-)} \leftrightarrow v_{(t^+)}$ is not in π , where $[t^-, t^+]$ is the interval of $RI_\Sigma(v)$ such that $t^- < t \leq t^+$; and this inclusion is strict, because $\pi \supset \pi_\gamma$. Therefore $Ch(\tau') \subset \gamma$, which contradicts the fact that γ explains minimally Σ .

- Let us show that $\gamma \in Cov(\Gamma(\pi_\gamma))$. Let us consider $\langle v_0, t_0 \rangle$ in γ , it means that $\tau_\gamma(t_0 - 1) \models \neg v_0$ and $\tau_\gamma(t_0) \models v_0$. We are going to show that it exists t^-, t^+ such that $t^- < t_0 \leq t^+$ and $\langle v_0, t^-, t^+ \rangle \in \Gamma(\pi_\gamma)$.
 - If it does not exist t^-, t^+ such that $t^- < t_0 \leq t^+$ and $[t^-, t^+] \in RI_\Sigma(v_0)$ then either v_0 does not appear in Σ or it appears only at one time point t^* . Let τ' be a trajectory satisfying Σ but in which there is no change for v_0 at time point t_0 this trajectory exists: $\forall v \in \mathcal{V} \setminus \{v_0\}, \forall t \neq t_0, \tau'(t)(v) = \tau_\gamma(t)(v)$ and $\forall t, \tau'(t)(v_0) = x$ where x is any truth value if v_0 never appears in Σ or is such that $\Phi_\Sigma(t^*) \models x$. It means that τ_γ is not a \preceq_{csi} -preferred trajectory for Σ since $Ch(\tau') = Ch(\tau_\gamma) \setminus \{\langle v_0, t_0 \rangle\}$. Therefore, there exists t^-, t^+ such that $t^- < t_0 \leq t^+$ and $[t^-, t^+] \in RI_\Sigma(v_0)$, and by definition of RI_Σ this interval $[t^-, t^+]$ is unique. Hence by definition of P_Σ and π_γ , we have $v_{0(t^-)} \leftrightarrow v_{0(t^+)} \in P_\Sigma \setminus \pi_\gamma$.
 - Now, we have to show that $\pi_\gamma \cup \{\Phi_\Sigma\} \models v_{0(t^+)}$. Suppose that it exists a trajectory τ satisfying $\pi_\gamma \cup \{\Phi_\Sigma\} \wedge \neg v_{0(t^+)}$. Since π_γ is a maximal consistent set, it means that τ should satisfy $v_{0(t^-)}$. Let us build a trajectory τ' similarly as above: namely,
 - * for each $v \in \mathcal{V}$ and each $t \in R_\Sigma(v)$, $\tau'(t)(v) = \tau(t)(v)$;
 - * for each $v \in \mathcal{V}$ and each $[t^-, t^+] \in RI_\Sigma(v)$ such that $v_{(t^-)} \leftrightarrow v_{(t^+)}$ is in π_γ , $\tau'(t)(v) = \tau(t^-)(v) = \tau(t^+)(v)$;
 - * for each $v \in \mathcal{V}$ and each $[t^-, t^+] \in RI_\Sigma(v)$ such that $v_{(t^-)} \leftrightarrow v_{(t^+)}$ is not in π_γ , there exists a t^* such that $t^- < t^* \leq t^+$ and $\langle v, t^* \rangle$ or $\langle \neg v, t^* \rangle$ is in γ . Then let

$$\tau'(t)(v) = \begin{cases} \tau(t^-)(v) & \text{if } t < t^*, \\ \tau(t^+)(v) & \text{if } t \geq t^*. \end{cases}$$

Since τ' coincides with τ for each $v \in \mathcal{V}$ and each $t \in R_\Sigma(v)$, we have $\tau' \models \Sigma \wedge \pi_\gamma \wedge v_{0t^-} \wedge \neg v_{0t^+}$. By construction all the changes in $Ch(\tau')$ belong to γ . By combining τ' and τ_γ in a new trajectory $\tau'' = \tau' \rightarrow^{t^*} \tau_\gamma^{t^*+1}$, we have found a trajectory that satisfies Σ since $\forall t, \tau''(t) \models \Sigma(t)$, and such that $Ch(\tau'') = \gamma \setminus \{\langle v_0, t_0 \rangle\}$ which contradicts the fact that τ_γ was minimal w.r.t. \preceq_{csi} . Hence, $\pi_\gamma \cup \{\Phi_\Sigma\} \models v_{0(t^+)}$.

The previous paragraph shows that for all $\langle v_i, t_i \rangle \in \gamma$ it exists one and only one pair of time points (t^-, t^+) such that $t^- < t_i \leq t^+$ and $\langle v_i, t^-, t^+ \rangle \in \Gamma(\pi_\gamma)$. The same result can be obtain for any $\langle \neg v_i, t_i \rangle \in \gamma$. Hence, $\gamma \in Cov(\Gamma(\pi_\gamma))$. \square

Proof of Proposition 15. Corollary of Lemmas 14 and 13. \square

Note that it is not always the case that the compact explanations Γ computed according to the previous method are *maximally* compact, as seen on the following example: $\Sigma = \langle a, c, a \vee c, b, \neg a \vee b, \neg a \rangle$ yields the compact minimal explanations $\{\langle \neg a, 4, 5 \rangle\}$ and $\langle \neg a, 6, 6 \rangle\}$ which are not *maximally* compact since their union $\{\langle \neg a, 4, 6 \rangle\}$ is a mcmE. Obtaining the mcmEs for Σ would require a compactification procedure whose precise description is outside the scope of the paper.

Note that in this section, the preference relation was fixed to \preceq_{csi} . However, this method can be adapted in order to find minimal explanations for many other preference relations. For instance, \preceq_{nc} would need to select consistent persistence sets of maximum cardinality instead of consistent sets maximum w.r.t. inclusion.

Due to Proposition 15 (and its generalizability to other preference relations), finding a covering set of compact explanations can be seen as a logic-based abduction problem (where abducibles correspond to elementary changes), and can be computed using dedicated algorithms. Complexity results and tractable classes [14] can be used to find some tractable subclasses of belief extrapolation.

7. Extrapolation and update

The position of extrapolation with respect to revision is central, and was exposed in Section 4. Here we discuss the position of extrapolation with respect to belief update, while Section 8 will consider other classes of belief change operators.

An update operator is a function mapping a knowledge base K representing knowledge about a system in an initial state and a new piece of information φ (whose precise meaning is to be discussed further) a new knowledge base $K \diamond \varphi$ representing the system after this evolution [18,34]. The key property of belief update is Katsuno and Mendelzon's postulate

$$\mathbf{U8} \quad (K1 \vee K2) \diamond \varphi \leftrightarrow (K1 \diamond \varphi) \vee (K2 \diamond \varphi)$$

which tells that models of K are updated independently.

7.1. Extrapolation is not update

Since belief update only considers two time points and takes as input a pair $((K, \alpha))$ of formulae referring respectively to $t = 1$ and $t = 2$, the comparison with extrapolation is better done in the restricted case where (a) $N = 2$ and (b) the input consists in two formulae referring to $t = 1$ and $t = 2$, which is exactly the case of scenarios of length 2. To be more precise, let $\Sigma = \langle \Sigma(1), \Sigma(2) \rangle$ be a scenario of length 2. Extrapolating Σ and then projecting the result on time 2 can be rephrased this way: from a belief set $K = \Sigma(1)$ at time 1 and an observation $\alpha = \Sigma(2)$ at time 2, compute a completed belief set at time 2. At first glance, this may look similar to belief update. However, this is not the case. More precisely,

Proposition 16. *Assume that the language contains at least two propositional symbols, i.e., $|\mathcal{V}| \geq 2$. Let $N = 2$ and E be an extrapolation operator. Define the operator \odot by: $\forall \varphi, \psi \in \mathcal{L}, \varphi \odot \psi = E(\varphi_{(1)} \wedge \psi_{(2)})^{\downarrow 2}$. Then \odot is not an update operator.*

The main reason for this result is that as soon as the language contains at least two propositional symbols, the AGM postulates are inconsistent with U8 (see for instance [16]). We reformulate the proof for extrapolation as follows:

Proof. Assume that $|\mathcal{V}| \geq 2$, and assume \odot satisfies (U8). Let $m1, m2$ and $m3$ be three different interpretations in $2^{\mathcal{V}}$ (which is possible because $|2^{\mathcal{V}}| \geq 4$). Consider the scenario $\Sigma = \langle \text{form}(m1, m2), \text{form}(m1, m3) \rangle$ (to simplify notations, we omit the curly brackets), because $\text{PERS} \wedge \Sigma$ is consistent (in other terms, there is no evidence that a change occurred), by (E2), we get $E(\Sigma) \equiv \text{PERS} \wedge \Sigma = \langle \text{form}(m1), \text{form}(m1) \rangle$, which by definition of \odot entails

$$(a) \quad \text{form}(m1, m2) \odot \text{form}(m1, m3) = \text{form}(m1)$$

Now, \odot satisfies (U8); therefore,

$$(b) \quad \text{form}(m1, m2) \odot \text{form}(m1, m3) \equiv \text{form}(m1) \odot \text{form}(m1, m3) \vee \text{form}(m2) \odot \text{form}(m1, m3).$$

From (a) and (b) we get

$$(c) \quad \text{form}(m1) \odot \text{form}(m1, m3) \vee \text{form}(m2) \odot \text{form}(m1, m3) \equiv \text{form}(m1).$$

Now, by (E3), $E(\langle \text{form}(m2), \text{form}(m1, m3) \rangle)$ is satisfiable; thus so is $\text{form}(m2) \odot \text{form}(m1, m3)$, which together with (c) implies

$$(d) \quad \text{form}(m2) \odot \text{form}(m1, m3) \equiv \text{form}(m1).$$

Now, by a symmetrical chain of implications, exchanging $m1$ and $m3$ in (a)–(d), leads to $\text{form}(m2) \odot \text{form}(m1, m3) \equiv \text{form}(m3)$, which contradicts (d). Therefore \odot cannot satisfy (U8). \square

This simple result has significant consequences. It means that for reasoning about time-stamped observations on a changing world, belief update is not adequate (see the example in Introduction). The key point is postulate U8 which, by requiring that all models of the initial belief set be updated separately, forbids us inferring new beliefs about the past from later observations.

In belief update, the input α should rather be interpreted as the projection of the expected effects of some “explicit change”, or more precisely, the *expected* (not the *observed*) effect of the action (or event) “make α true”. See [20] for further discussion.

We see that the crucial issues are *observability* (what do we observe about the world at what time?) and *predictability of change*. Belief extrapolation deals with observation and unexpected change, while belief update is suitable for expected change without observations. In Sandewall's taxonomy [27], extrapolation is adequate for the action-free subclass of $\mathcal{K}\text{-IS}$ (correct knowledge, inertia and surprises) while update is adequate for the class $\mathcal{K}_p\text{-IA}$ (no observations after the initial time, inertia and alternative results of actions).

7.2. Non-inertial extrapolation and generalized update

Generalized update [7] consists in finding out which events (from a given set of events E) most likely occurred between two time points t_1 and t_2 , and use the knowledge about the dynamics of these events to reason about what was true at t_1 and what is true at t_2 . The plausibility of events, as well as their dynamics, are modeled by ordinal conditional functions. In Section 3.2, we describe an event-based extrapolation operator in the same manner as [7]. But in order to

impose inertia we impose the existence of the null event that was only “occasionally assumed” by Boutilier. For Boutilier, “The null event ensures (with certainty) that the world does not change” and is such that $\kappa(e_{\emptyset}|m) = 0$ if and only if $e = e_{\emptyset}$, and $\kappa(m'|m, e_{\emptyset}) = \infty$ if $m \neq m'$. We slightly modify the second condition into $\kappa(m'|m, e_{\emptyset}) = 0$ if and only if $m = m'$ in order to impose that static trajectories are normal. Given an event model \mathcal{E} , the plausibility of the transition from m to m' through e is $\kappa_{\mathcal{E}}(e, m, m') = \kappa(m) + \kappa(e|m) + \kappa(m'|m, e)$, and the plausibility of the transition from m to m' is defined as $\kappa_{\mathcal{E}}(m, m') = \min_{e \in E} \kappa_{\mathcal{E}}(e, m, m')$.

If we consider non-inertial extrapolation operator by dropping the inertia assumption, then generalized update (w.r.t. event model \mathcal{E}) is the restriction of non-inertial extrapolation to scenarios with two time points and the preference relation defined as follows: $\langle m_1, m_2 \rangle \preceq_{\mathcal{E}} \langle m_3, m_4 \rangle$ if $\kappa_{\mathcal{E}}(m_1, m_2) \leq \kappa_{\mathcal{E}}(m_3, m_4)$. The proof is omitted – it is rather straightforward.

Now, we know that generalized update generalizes both belief revision and belief update, which suggests that non-inertial belief extrapolation generalizes belief update. This is indeed true to some extent: let \diamond be a KM-update operator,¹¹ fix a propositional formula α , and consider the following event-based model $\mathcal{E}_{\diamond, \alpha}$ with only one event e_{α} , $\kappa(e_{\alpha}|m) = 0$ for every m , and $\kappa(m'|m, e_{\alpha}) = 0$ if and only if $m' \in m \diamond \alpha$, that is, if and only if $m' \in \text{Pref}(\preceq_m, \text{Mod}(\alpha))$, where $\{\preceq_m, m \in \mathcal{M}\}$ is the collection of faithful preorderings associated with \diamond . e_{α} is an event “simulating” an update by α , and e_{α} being the only event in the event model implies that this update occurs for sure. Call $\preceq_{[\diamond, \alpha]}$ the preference relation associated with this event model. Then $\text{Ex}_{[\diamond, \alpha]}(\langle \varphi, \top \rangle) = \langle \varphi, \varphi \diamond \alpha \rangle$. (This holds because for any 2-trajectory τ , $\kappa_{\mathcal{E}_{\diamond, \alpha}}(\tau) = 0$ holds if and only if $\tau(1) \models \varphi$ and $\tau(2) \models \tau(1) \diamond \alpha$.) Therefore, to some extent, update is a particular case of non-inertial extrapolation. This is not in contradiction with Proposition 16, however, for the following two reasons: (a) inertia (or persistence), needed in Proposition 16, and (b) the way used for defining the belief extrapolation operator is different: in Proposition 16, we define one extrapolation operator, of which φ and α are arguments; whereas here we use a family of extrapolation operators – actually a huge family, since, even when \diamond is fixed, we have a separate operator for each α . Therefore, rather than saying that non-inertial extrapolation generalizes belief update, we should rather say that for a fixed formula α , there exists a non-inertial extrapolation operator simulating an update by α .

7.3. Integrating extrapolation and update

Extrapolation and update complete each other and, in order to be able to reason both with implicit and explicit change, we can integrate both, as suggested in Section 5.2 of [13]. This is developed in [12] and applied to the study of causal ascription in a scenario. This approach defines a kind of hypothetical reasoning using both update and extrapolation, which amounts to compute what could have happened if something had been different in a given story. This operation consists in updating (in the KM sense) a scenario by a temporal formula: given an initial scenario Σ , it consists in computing, for each trajectory τ satisfying Σ , the closest trajectories to τ satisfying the new temporal formula Ψ .

Such an integration of unexpected change (via belief extrapolation) and ontic actions (via belief update) has been investigated in a few other proposals that we describe in Section 8.

8. Related work

Belief extrapolation is not the only class of belief change operators dealing with unexpected change. We review below a series of approaches, which all have in common that they minimize unexpected change. Some of them are very specific operators, some others are more general. Some of them also consider other features such as actions, or consider other ways of explaining a discrepancy between the agent's expected beliefs and the actual state of affairs, such as fallible observations, or wrong initial beliefs about the world.

8.1. Berger, Lehmann and Schlechta

Probably the most related approach to belief extrapolation is the generic class of belief change operators proposed by Berger et al. [5]. This class of belief change operators is actually a subclass of the set of extrapolation operators (although the authors use the terminology “iterated updates” – which we think is not adequate, as discussed in Section 7), and is probably the first approach, chronologically speaking, on belief extrapolation. More exactly, an extrapolation operator is a belief change operator as in [5] if and only if its preference relation satisfies the following property (BLS). Let us first define $h(\tau)$ by $h(\tau)(1) = \tau(1)$ and for $k \geq 1$, $h(\tau)(k+1) = \tau(i_{k+1})$ where i_{k+1} is the smallest integer such that $\tau(i_{k+1}) \neq \tau(i_k)$ if it exists, and $i_1 = 1$; and so on until such an integer can no longer be found. Intuitively, the trajectory $h(\tau)$ is a shortcut for the trajectory τ in which the persistent periods are removed: only the time points where the trajectory has evolved are kept.

BLS property For any trajectories τ, τ' such that there exists a sequence of integers $1 \leq j_0 < j_1 < \dots < j_k$ such that $h(\tau) = \langle m^0, \dots, m^n \rangle$ and $h(\tau') = \langle m^{j_0}, \dots, m^{j_k} \rangle$, we have $\tau < \tau'$.

¹¹ The argument would hold more generally with any operator satisfying (U8), even if some of the other postulates are not satisfied.

It can be shown that, among the preference relations we proposed, those that satisfy (BLS) are \preceq_{nc} , \preceq_{ncl} , \preceq_{ifc} , \preceq_{cseq} and \preceq_k . While [5] focus on the axiomatic properties of operators and representation results, they do not study various preference relations on trajectories nor investigate computational issues.

8.2. Hunter and Delgrande

The framework developed in [17] allows not only for unexpected changes, but also fallible observations, and actions (without exceptional effects). Their approach is somehow similar to extrapolation, in the sense that it is based on the selection of preferred trajectories. On the one hand, they go much beyond extrapolation, since the latter does not deal with ontic action effects, nor fallible observations. On the other hand, they commit to a specific choice of a preference relation, that makes use of integer-valued ranking functions that can be seen as associating “surprise degrees” both to unexpected events and incorrect observations (the reason for using integer-valued degrees is probably that it makes easier the combination of “surprise degrees” of different kinds). This restriction on preference relations is the main reason for their Proposition 7, that states that some extrapolation operators are not representable in their framework.¹²

8.3. Shapiro and Pagnucco

Shapiro and Pagnucco [30] introduce an account of reasoning about action and change in the situation calculus in which sensing that does not accord with beliefs can be resolved in one of two ways: deciding that a belief about the initial situation was mistaken, or hypothesizing the occurrence of a sequence of exogenous actions. They hypothesize the occurrence of exogenous actions only when necessary, and do so through a notion of minimal change. Therefore, the agent will prefer situations that have less exogenous actions.

Apart from the fact that it also allows ontic actions (and thus updates), this framework is a situation calculus version of a specific extrapolation operator (minimizing the number of exogenous events). Our other extrapolation operators could probably be expressed in the situation calculus as well (we leave this for further research).

8.4. Booth and Nittka

Booth and Nittka [6] take a subjective view of belief revision that could be seen as a subjective version of extrapolation: given an observer and an observed agent, the observer tries to make inferences about what the agent believed (or will believe) at a given moment, based on an observation of how the agent has responded to some sequence of previous belief revision inputs over time. Assuming a framework for iterated belief revision which is based on sequences, they construct a model of the agent that “best explains” the observation.

As they say, a fundamental difference between that work and extrapolation is that belief extrapolation is, like traditional operators of revision and update, an agent’s perspective operator: it is concerned with how an agent should form a picture of how the external world is evolving, whereas they are interested in forming a picture of how an observed agent’s beliefs are evolving.

The comparison between [6] and extrapolation suggests a promising issue for further research: “subjective extrapolation” would consist in starting with a scenario describing what we know of the agent’s beliefs at different time points (possibly using formulae from doxastic logic), and then find the most plausible events that occurred and that explain the changes in the agent’s beliefs.

8.5. Liberatore and Schaerf

Liberatore and Schaerf [23] propose a fairly general system (BRELS) aiming at integrating revision, update and merging. It deals with time-stamped observations and consider two semantics; using the “trajectory” semantics and assuming that there is no more than one observation at each time point, we obtain our extrapolation operator induced by the penalty-induced relation \preceq_k . The other semantics (“pointwise”) yields iterated update (but is incompatible with extrapolation because of U8 which underlies this semantics). On the other hand, BRELS provides an integrated treatment of both static and dynamic reasoning, and thus includes the possibility of representing exceptional effects or misperceptions.

8.6. Friedman and Halpern

Friedman and Halpern [15] define a very general framework for belief change, of which revision and update are two specific instances. We give here a simplified presentation of the framework (and slightly change the terminology), but rich enough to show how extrapolation fits in it. Let $W = 2^{P^S}$ the set of all truth assignments to primitive propositions. A *plausibility measure* is a mapping Pl from 2^W to an ordered scale D with a top element \top_D and a bottom element \perp_D ,

¹² When specializing the framework so that observations are fully reliable and no ontic actions are involved, the class of extrapolation operator obtained would be a specific subclass of event-based extrapolation operators.

and such that $Pl(W) = \top_D$, $Pl(\emptyset) = \perp_D$ and $A \subseteq B$ entails $Pl(A) \leq_D Pl(B)$. A *belief change system* (see Section 4 of [15]) is a pair $I = (\mathcal{R}, \mathcal{P})$ where

- \mathcal{R} is a set of runs; a run r is a pair $\langle \tau_r, \Sigma_r \rangle$ where τ_r is a trajectory and Σ_r is a scenario. If r is a run and t a time point we denote $r_e(t) = \tau_r(t)$ and $r_a(t) = \langle \Sigma_r(1), \dots, \Sigma_r(t) \rangle$. $r_e(t)$ (resp. $r_a(t)$) is the *environment state* (resp. the *agent's local state*) at (r, t) .
- \mathcal{P} is a plausibility assignment mapping each pair (r, t) to a plausibility measure $Pl_{(r,t)}$ on 2^W .

A belief change is said to satisfy the **PRIOR** assumption if at the start of the process, the agent has a *prior plausibility* $Pl_a: \mathcal{R} \rightarrow D$ on the set of runs \mathcal{R} from which all $Pl_{(r,t)}$ can be determined. At time point t_0 of the run r , the set of runs considered possible for the agent is the set of all runs such that (a) the observations done so far, i.e. $\Sigma(t)$ from $t = 1 \dots t_0$, coincide with those in r_a : $\Sigma(t) = \Sigma_r(t)$; (b) the objective part of the run is compatible with the observations done so far, i.e., for all $t = 1 \dots t_0$, $\tau_r(t) \models \Sigma(t)$. Under the key assumption named **PRIOR** (that we do not recall here), belief change is made *conditioning*: at each time point, the agent modifies its current plausibility distribution on runs by conditioning it by whatever piece of information he has learned. Finally, a formula φ is believed by the agent in some point (r, t) if for all $s \in W_{(r,t)}$ such that there is no $s' \in W_{(r,t)}$ with $Pl(s') > Pl(s)$, then $\pi(s) \models \varphi$.

Belief revision and update are special instances of belief change systems satisfying **PRIOR**. Now, for all classes of preference relations studied in this article, scenario extrapolation is also an instance of a belief change system satisfying *prior*. The key point consists in identifying plausibility spaces and conditioning for various extrapolation operators. Intuitively, the prior plausibility at the starting point of the process (at time 0, that is, before the first observation is made) reflects the preference relation over trajectories. To make things simpler, we illustrate the correspondence on the case $\preceq = \preceq_{nc}$, but we insist that the process works for a much wider class of extrapolation operators (including all those that are based on any of the preference relations given in Section 3).

Let $D = \overline{\mathbb{N}}$, with the reverse natural order, namely $i \leq_D j$ iff $i \geq j$, and $\top_D = 0$, $\perp_D = +\infty$. For any run r , define $pl(r) = |Ch(\tau_r)|$: the plausibility of a run is the number of changes of its associated trajectory. Conditioning by observations is defined by

$$pl(r \mid \varphi_{(t)}) = \begin{cases} +\infty & \text{if } \tau \models \neg\varphi_{(t)}, \\ pl(\tau_r) - \min\{pl(\tau') \mid \tau' \models \varphi_{(t)}\} & \text{otherwise.} \end{cases}$$

Since $pl(r)$ is fully determined by τ_r , we use the notation $pl(\tau_r) = pl(r)$ (that is, the plausibility of a trajectory is defined as the plausibility of any run that corresponds to it). The plausibility of a set X of runs is $Pl(X) = \min\{pl(r) \mid r \in X\}$: it is thus determined by the number of changes in the trajectory of the most static run in X . The definition for $Pl(X \mid \varphi_{(t)})$ is similar. Finally, let $run(\Sigma) = \{r \mid \tau_r \models \Sigma\}$. Now, we have the following:

$$Ex(\Sigma) \models \Sigma' \quad \text{iff} \quad \text{for any } \tau, \quad pl(\tau \mid \Sigma) = 0 \text{ implies } \tau \models \Sigma'$$

(the proof of which is omitted).

Example 10. For example, let $\Sigma = \langle a, \neg a \wedge b, a \vee \neg b, a \wedge b \rangle$. Initially, $pl_0(\tau) = |Ch(\tau)|$. After observing $a_{(1)}$, we get

$$pl_1(\tau) = pl_0(\tau \mid a_{(1)}) = \begin{cases} |Ch(\tau)| & \text{if } \tau(1) \models a, \\ +\infty & \text{otherwise.} \end{cases}$$

Then, at time 2,

$$\begin{aligned} pl_2(\tau) &= pl_1(\tau \mid \neg a_{(2)} \wedge b_{(2)}) \\ &= pl_0(\tau \mid a_{(1)} \wedge \neg a_{(2)} \wedge b_{(2)}) \\ &= \begin{cases} |Ch(\tau)| - 1 & \text{if } \tau \models a_{(1)} \wedge \neg a_{(2)} \wedge b_{(2)}, \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

We have $pl_2(\tau) = 0$ if and only if $\tau(1) = ab$ and $\tau(2) = \neg ab$. At time 4, the two trajectories such that $pl_4(\tau) = 0$ are $\langle ab, \neg ab, \neg ab, ab \rangle$ and $\langle ab, \neg ab, ab, ab \rangle$. The trajectories τ such that $pl(\tau \mid \Sigma) = 0$ are exactly the trajectories minimizing the total number of changes, i.e., the preferred trajectories with respect to \preceq_{nc} .

Characterizing other specific extrapolation operators considered in this paper by a conditioning operator can be done in a very similar way.

8.7. Extrapolation and dynamic diagnosis

Belief extrapolation addresses a problem similar to dynamic diagnosis, where the goal is to identify a complete evolution of the world which best fits a given (maybe incomplete) history that contains some abnormal behaviours. More precisely, the general principle of dynamic diagnosis is composed of the following elements:

- a system, defined by its components and their different behaviour modes (including the normal behaviour and possibly several failure modes);
- a set of actions, including test actions (such as measures) and repair actions, and a series of known action occurrences at different time points;
- a series of observations made at different time points.

We then look for a set of plausible diagnoses, that is, a series of faults together with the time where they occurred, or for a plan for having the system work again normally. There are various approaches of dynamic diagnosis (see e.g. [8] for an early survey of definitions) which vary along with several features, such as the logical language used for modelling the system, the actions and the observations (e.g., action languages in [32] and [3], logic programs in [2]), the nature of the available actions, the type of diagnosis (abductive or consistency-based), and the selection principle for finding plausible diagnoses. Consistency-based dynamic diagnosis consists in finding a series of failures across time which is *consistent* with the observations and the system description (while an abductive diagnosis together with the system description should allow to *deduce* the observations). Therefore, extrapolation can be seen both as a simplification and a generalization of consistency-based dynamic diagnosis: a simplification, because extrapolation assumes that the normal behaviour of the system is inertia and that no action is available to the agent; and a generalization, because extrapolation proposes a richer set of selection operators (which not only minimize the set or number of changes but also are able to compare sequences of changes, combinations of changes etc.), while each individual approach to dynamic diagnosis makes use of a specific selection criterion. For instance, the criterion in Thielscher [32] is chronological and minimizes abnormalities in the initial state while taking account of the prior likelihood of failure of the components; Baral et al. [3] minimize the set of faulty components and exogenous actions and propose a diagnosis-plan that aims at selecting dynamically a diagnosis by giving the sequence of tests to be done in order to discriminate between the unobserved possible faulty components; this kind of checking plan is also used by Balduccini and Gelfond [2] who do not use any prior minimization process but filter out some of the candidate diagnoses by using some elimination criteria (such as being unrelated to the symptoms, a too large time interval between faults and symptoms, or a too large number of faults).

Thus, extrapolation offers a much more systematic and principled study of such operators (through an axiomatization and a representation result), and our results are of course relevant to dynamic diagnosis, especially because extrapolation contributes to bridging dynamic diagnosis and belief change.

9. Conclusion

We have studied a general family of belief change operators which consist in completing initial observations by persistence assumptions. We have discussed in details the links between extrapolation and revision, and its differences with belief update. Belief extrapolation amounts to a specific case of belief revision where beliefs are expressed in a time-stamped language, and possesses many specificities that are proper to this temporal structure. Belief extrapolation can also be seen as a simplification of dynamic diagnosis, and thus contributes to build a bridge between dynamic diagnosis and belief change.

The connection between belief revision and belief extrapolation deserves some final comments. As we have seen in Section 4, every belief extrapolation operator can be expressed as a belief revision operator through time-stamping the propositional language. Now, knowing that belief extrapolation is, in some sense, less general than belief revision, does not mean it does not deserve a study of its own, and most of our results concern specificities of extrapolation that would be meaningless in a standard, atemporal, belief revision framework. This is the case for the preference relations and extrapolation operators described in Section 3, together with their associated complexity results in Section 6: none of these could be defined in an atemporal language (to begin with, change sets would be meaningless). Section 5 lays the focus on a few specific temporal properties of extrapolation. Finally, the method given in Section 6 also exploits the temporal structure as much as possible.

We chose to model beliefs across time using explicit time points (thus making use of a propositional logic of reified time). Since our results, and the preference relations we have focused on, do not exploit the metric nature of time, they would still hold in a similar way if we had chosen to use a purely symbolic representation of time such as in modal logics for belief change (e.g., [29,33]) or epistemic or doxastic extensions of the situation calculus (e.g., [11,28]). In such frameworks we are still able to express sequences of observations and reason about them. More generally, given that belief extrapolation is a form of belief revision, all frameworks that have been shown relevant to belief revision are also relevant to belief extrapolation. Finally, using a modal language would open the door to new opportunities; in particular, having a specific belief modality for each time point (or situation) allows to express mutual intertemporal beliefs such as “at time 3 I believed that φ held at time 2 while I now believe that φ did not hold at time 2”, as in the following example: let $\Sigma = \langle a \wedge b \wedge c, (\neg a \wedge \neg b) \vee \neg c, \top, \neg a, \neg b \rangle$. Performing an Ex_{nc} extrapolation after only three observations have been made, we extrapolate that $\neg c$ held at time 2, while extrapolating after all observations gives that c held at 2.

Further work includes developing a full framework for reasoning about both predicted change (action effects) and unpredicted change, in the same vein as some of the works reported below, but with a much larger degree of generality. Moreover, in this paper we focused on *inertial* extrapolation, which considers changes as exceptional (which, technically, leads to minimizing change). We may consider more general extrapolation operators that deal with non-inertial systems with a specific dynamics. Another line of further work consists in generalizing extrapolation to nontemporal classes of sce-

narios, such as spatial scenarios, where observations are labelled with points in space and persistence by default applies to adjacent points. See [10] for a preliminary study.

Acknowledgements

We would like to thank the anonymous reviewers for their very useful comments, as well as Richard Booth, Marie-Odile Cordier, Jim Delgrande, Didier Dubois, Andreas Herzig, Sébastien Konieczny, Pierre Marquis, Alexander Nittka, Odile Papini and Torsten Schaub for helpful discussions.

References

- [1] C. Alchourrón, P. Gärdenfors, D. Makinson, On the logic of theory change: partial meet contraction and revision functions, *Journal of Symbolic Logic* 50 (1985) 510–530.
- [2] M. Balduccini, M. Gelfond, Diagnostic reasoning with A-Prolog, *Theory and Practice of Logic Programming* 3 (2003) 425–461.
- [3] C. Baral, S. McIlraith, T. Cao Son, Formulating diagnostic problem solving using an action language with narratives and sensing, in: *Proceedings of KR2000, Breckenridge, Colorado (USA), 2000*, pp. 311–322.
- [4] S. Benferhat, S. Konieczny, O. Papini, R. Pino-Pérez, Iterated revision by epistemic states: axioms, semantics and syntax, in: *Proceedings of ECAI-2000, 2000*, pp. 13–17.
- [5] S. Berger, D. Lehmann, K. Schlechta, Preferred history semantics for iterated updates, *Journal of Logic and Computation* 9 (6) (1999) 817–833.
- [6] R. Booth, A. Nittka, Reconstructing an agent's epistemic state from observations, in: *Proceedings of IJCAI-05, 2005*, pp. 394–399.
- [7] C. Boutilier, A unified model of qualitative belief change: a dynamical systems perspective, *Artificial Intelligence* 1–2 (1998) 281–316.
- [8] V. Brusoni, L. Console, P. Terenziani, D. Theseider Dupré, A spectrum of definitions for temporal model-based diagnosis, *Artificial Intelligence* 102 (1) (1998) 39–79.
- [9] A. Darwiche, J. Pearl, On the logic of iterated revision, *Artificial Intelligence* 89 (1997) 1–29.
- [10] J. Delgrande, J. Lang, T. Schaub, Belief change based on global minimisation, in: *Proceedings of IJCAI-07, 2007*, pp. 2468–2473.
- [11] R. Demolombe, M. Pilar Pozos-Parra, A simple and tractable extension of situation calculus to epistemic logic, in: *Proceedings of the Twelfth International Symposium on Intelligent Systems*, in: *LNAI*, vol. 1032, 2000, pp. 515–524.
- [12] F. Dupin de Saint-Cyr, Scenario update applied to causal reasoning, in: *Proceedings of KR2008, 2008*, pp. 188–197.
- [13] F. Dupin de Saint-Cyr, J. Lang, Belief extrapolation (or how to reason about observations and unpredicted change), in: *Proceedings of KR2002, 2002*, pp. 497–508.
- [14] T. Eiter, G. Gottlob, The complexity of logic-based abduction, *Journal of the Association for Computing Machinery* 42 (1) (1995) 3–42.
- [15] N. Friedman, J.Y. Halpern, Modelling beliefs in dynamic systems. Part II: revision and update, *Journal of Artificial Intelligence and Research* 10 (1999) 117–167.
- [16] A. Herzig, Logics for belief base updating, in: *Handbook of Defeasible Reasoning and Uncertainty Management*, vol. 3, Kluwer Academic Publishers, 1998, pp. 189–231.
- [17] A. Hunter, J.P. Delgrande, Belief change in the context of fallible actions and observations, in: *Proceedings of AAAI'06, 2006*.
- [18] H. Katsuno, A.O. Mendelzon, On the difference between updating a knowledge base and revising it, in: *Proceedings of KR'91, 1991*, pp. 387–394.
- [19] H. Katsuno, A.O. Mendelzon, Propositional knowledge base revision and minimal change, *Artificial Intelligence* 52 (1991) 263–294.
- [20] J. Lang, Belief update revisited, in: *Proceedings of IJCAI-07, 2007*, pp. 2517–2522.
- [21] J. Lang, P. Liberatore, P. Marquis, Propositional independence – formula-variable independence and forgetting, *Journal of Artificial Intelligence Research* 18 (2003) 391–443.
- [22] D. Lehmann, Belief revision, revised, in: *Proceedings of IJCAI-95, 1995*, pp. 1534–1540.
- [23] P. Liberatore, M. Schaerf, BRELS: A system for the integration of knowledge bases, in: *Proceedings of KR2000, 2000*, pp. 145–152.
- [24] F. Lin, R. Reiter, Forget it! in: *Proceedings of the AAAI Fall Symposium on Relevance, New Orleans, 1994*, pp. 154–159.
- [25] B. Nebel, How hard is it to revise a belief base?, in: D. Dubois, H. Prade (Eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, Kluwer Academic, 1998, pp. 77–145.
- [26] C.H. Papadimitriou, *Computational Complexity*, Addison-Wesley, 1994.
- [27] E. Sandewall, *Features and Fluents*, Oxford University Press, 1995.
- [28] R. Scherl, H. Levesque, Knowledge, action, and the frame problem, *Artificial Intelligence* 144 (1–2) (2003) 1–39.
- [29] K. Segerberg, Belief revision from the point of view of doxastic logic, *Bulletin of the IGPL* (1995) 535–553.
- [30] S. Shapiro, M. Pagnucco, Iterated belief change and exogeneous actions in the situation calculus, in: *Proceedings of ECAI'04, 2004*, pp. 878–882.
- [31] Y. Shoham, *Reasoning about Change – Time and Causation from the Standpoint of Artificial Intelligence*, MIT Press, 1988.
- [32] M. Thielscher, A theory of dynamic diagnosis. *Electronic transactions on artificial intelligence*, *Electronic Transactions on Artificial Intelligence* 1 (1997) 73–104.
- [33] J. van Benthem, Dynamic logic for belief revision, *Journal of Applied Non-Classical Logics* 2 (17) (2007) 129–155.
- [34] M. Winslett, Reasoning about action using a possible models approach, in: *Proceedings of AAAI'88, St. Paul, 1988*, pp. 89–93.