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

Florence Dupin de Saint-Cyr
LERIA,
Université d'Angers (France)
bannay@info.univ-angers.fr

Jérôme Lang
IRIT,
Université Paul Sabatier, Toulouse (France)
lang@irit.fr

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 semantic and we propose several extrapolation operators. We study some properties verified by these operators and investigate computational issues, including computational complexity. We study in detail the position of belief extrapolation with respect to revision and update ; in particular, we show that, although it deals with non-static worlds, belief extrapolation cannot be a belief update operator in the sense of Katsuno and Mendelzon; then, we show how to integrate update and extrapolation. Lastly, closely related lines of work are positioned with respect to belief extrapolation.

1 INTRODUCTION

The distinction between belief revision and update [12] is usually thought of this way: revision incorporates newly acquired beliefs about a static world while update modifies beliefs according to a change in the system, "notified" by an "input" formula φ . However, there is some ambiguity in the exact nature of the input formula φ . What does "a notification of the change" mean exactly? At least two distinct interpretations are possible:

- some time has passed between the initial time point t_0 when the initial belief state K was known to hold and a later time point t_1 when φ is *observed* (by means of a sensor, a request to a database, a notification by another agent, etc.);
- φ is the *expected consequence* of an action performed between t_0 and t_1.

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

We consider a system to diagnose, whose components may fail independently, which means that there is no prior causal link between the failure of a component and the failure of another one. To simplify the problem we assume that for a given component, the transition from the failure mode to the working mode is not more exceptional than the transition from the working mode to the failure mode[1].

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

Scenario 2: *at $t = 1$, as for Scenario 1, $K = (faulty\_a \oplus faulty\_b) \wedge \neg faulty\_c$. Then the action of repairing component b as well as the (unfortunate) action of breaking component c are performed. The effects of these two independent actions can be expressed by $\varphi = \neg faulty\_b \wedge faulty\_c$.*

These two scenarios are different and must lead to different conclusions. In case of Scenario 1, the intuitive way of changing beliefs is: "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; as to c that has just been oberved faulty, there is no am-

[1] This assumption is not so irrealistic, considering that other agents can repair a faulty component without the "main" agent (the beliefs of whom we are interested in) being aware.

biguity". Therefore, the belief state at time 2 should be $faulty\_a \wedge \neg faulty\_b \wedge faulty\_c$. The case of Scenario 2 is different: the effect of repairing 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 $\neg faulty\_b \wedge faulty\_c$.

Now, belief update may be the right thing to do for Scenario 2, but not for Scenario 1. The principle responsible to this inadequacy of belief update to handle Scenario 1 is the most "typical" property of belief update, namely U8 (see Section 5) which says that models of K should be updated independently; as a consequence, usual belief update operators satisfy $K \diamond \varphi = \neg faulty\_b \wedge faulty\_c$; this is the expected result for Scenario 2 but not for Scenario 1 (where we expect $faulty\_a \wedge \neg faulty\_b \wedge faulty\_c$).

From these two scenarios we may advance an hypothesis: *belief update is not appropriate when the input is an observation.* Two questions follow: (Q1) When is it appropriate? (Q2) How should we change a belief base to take account of observations occuring at different time points? As to Q1, 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.* Q2 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. Such operators are not completely new (they have been considered in [3] and [17] – see the related work Section). The basic assumptions for extrapolation operators are the following two ones:

1. *the agent can only observe some properties at some time points*, but does not have the ability of performing actions; furthermore, if events occur, they are perceived by the agent through the observations only (for instance, the occurrence of the event that it will be raining tonight will be perceived by actually *seeing* the rain, or seeing the wet ground tomorrow morning, but not by listening to today's weather forecast). This is why changes can be qualified as *unpredicted.* In Section 5, however, we show that it is possible to generalize extrapolation operators so as to handle actions as well, provided that they are distinguished from observations, in order to minimize only the fluents not concerned by action affects.

2. *the system is inertial*, which means that by default, the system remains in a static state. This assumption justifies the use of a change minimization policy.

The rest of the paper is organized as follows. In Section 2 we give some background about revision and update. In Section 3 we formally define belief extrapolation; we propose several extrapolation operators and study their properties. In Section 4 we investigate computational issues, and propose a practical method for computing extrapolated beliefs for one of the formerly considered extrapolation operators. In Section 5 and 6, we study the connections between extrapolation and, respectively, revision (including iterated revision) and update. Namely, we show in Section 4 that under reasonable assumptions, extrapolation can be seen as a particular belief revision process where we revise the beliefs about the persistence of fluents by the time-stamped observations, and in Section 5 that a belief extrapolation operator cannot be a belief update operator unless very strong assumptions are made; rather, we show that extrapolation and update are complementary and that a full framework for reasoning about beliefs across time should involve both of them. This helps us assessing the respective positions of belief update and belief extrapolation in Sandewall's taxonomy of logically represented dynamical systems. Connections with several other works, including pioneering works on extrapolation [3] and [17], are investigated in Section 6.

2 BELIEF EXTRAPOLATION

2.1 SCENARIOS AND TRAJECTORIES

Let $\mathcal{L}$ be a propositional language built from a finite set of propositional variables $\mathcal{V} = \{v_1, ..., v_n\}$, the usual connectives, and the Boolean constants $\top$ (tautology) and $\bot$ (contradiction). $\mathcal{M} = 2^{\mathcal{V}}$ is the set of interpretations for $\mathcal{V}$. Formulas of $\mathcal{L}$ are denoted by Greek letters (φ, ψ etc.) and interpretations are denoted by m_i, m_j etc. If $\varphi \in \mathcal{L}$ then $Mod(\varphi)$ is the set of all models of φ. A *fluent* f is a literal, i.e., a variable or its negation; if $f = v_i$ (resp. $\neg v_i$) then $\neg f$ is $\neg v_i$ (resp. v_i). The set of fluents is $\mathcal{F} = \{v_1, \neg v_1, ..., v_n, \neg v_n\}$.

Definition 1 (scenarios)
A scenario $\Sigma = \langle \varphi_1, ..., \varphi_N \rangle$ *is a finite sequence of propositional formulas. We denote by* $\Sigma(t)$ *the* t^{th} *element* φ_t *of the sequence.* $|\Sigma|$ *denotes the length of* Σ. $\mathcal{S}$ *is the set of all scenarios.*
Σ *is said to be* consistent *if and only if for each* $t \in 1..|\Sigma|$, $\Sigma(t)$ *is consistent.*

For any two scenarios Σ, Σ' *of length* N, *we define:*
$\Sigma \models \Sigma'$ *if and only if* $\forall t$, $\Sigma(t) \models \Sigma'(t)$;

$\Sigma \equiv \Sigma'$ *if and only if* $\Sigma \models \Sigma'$ *and* $\Sigma' \models \Sigma$;
$\Sigma \wedge \Sigma' = \langle \Sigma(t) \wedge \Sigma'(t) | t \in 1 \ldots N \rangle$.

The concatenation $\Sigma.\Sigma'$ *of two scenarios* Σ *and* Σ', *is defined by:*

$$(\Sigma.\Sigma')(t) = \begin{cases} \Sigma(t) & \text{if } t \leq |\Sigma| \\ \Sigma'(t - |\Sigma|) & \text{if } t > |\Sigma| \end{cases}$$

The scenario Σ *s.t.* $\Sigma(t) = \varphi$ *and* $\Sigma(t') = \top$ *for all* $t' \neq t$ *is denoted by* $[t]\varphi$.

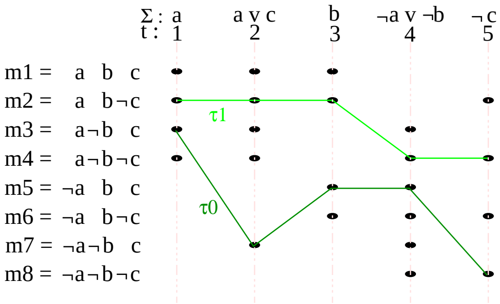
Definition 2 (trajectories)
A trajectory *of length* N *is a sequence* $\tau = \langle \tau(1), .., \tau(N) \rangle$ *of interpretations (*i.e., $\tau(i) \in \mathcal{M}$*).* $TRAJ_N$ *denotes the set of all trajectories of length* N.

If Σ *is a scenario and* τ *a trajectory of the same length, we say that* τ satisfies Σ *(denoted by* $\tau \models \Sigma$*) if and only if* $\forall t \in 1..N$ *we have* $\tau(t) \models \Sigma(t)$.

$Traj(\Sigma) = \{\tau \in TRAJ_{|\Sigma|} | \tau \models \Sigma\}$ *is the set of trajectories satisfying* Σ.

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

Example 1 *Let* $\Sigma = \langle a, a \vee c, b, \neg a \vee \neg b, \neg c \rangle$ *and the possible interpretations* $m1, \ldots, m8$ *as on the figure below. Many trajectories satisfy* Σ, *for instance* $\tau_0 = \langle m3, m7, m5, m5, m8 \rangle$ *and* $\tau_1 = \langle m2, m2, m2, m4, m4 \rangle$.



Note that there is no static trajectory satisfying Σ.

Definition 3 (change set)
The change set $Ch(\tau)$ *of a trajectory* τ *is defined by:*

$$Ch(\tau) = \left\{ \langle f, t \rangle \;\middle|\; \begin{array}{l} f \in \mathcal{F},\ t \in 2..|\tau|, \\ \tau(t-1) \models \neg f \text{ and } \tau(t) \models f \end{array} \right\}$$

We also define $Ch(\tau, f) = \{t \mid \langle f, t \rangle \in Ch(\tau)\}$.

Example 1, continued
$Ch(\tau_0) = \{\langle \neg a, 2 \rangle, \langle b, 3 \rangle, \langle \neg b, 5 \rangle, \langle \neg c, 5 \rangle\}$ *and*
$Ch(\tau_1) = \{\langle \neg b, 4 \rangle\}$.
We see that τ_1 *looks "more static" than* τ_0.

The input of an extrapolation operator is a consistent scenario Σ. Formulas of Σ are *observations* (coming from sensors, from communication with other agents etc.). For the sake of simplicity we do not consider static laws expressing dependence between fluents; however, it is possible (though not optimal in practice) to give these static laws the status of observations: at each time point the initial belief will thus be the conjunction of the observation and the static laws.

The consistency of Σ means that observations are assumed to be reliable (see [10] for a discussion on this point).

2.2 PREFERRED TRAJECTORIES AND EXTRAPOLATION OPERATORS

Given an observation scenario Σ, *belief extrapolation* consists in completing it using persistence assumptions (such a process is referred to as *chronicle completion* in [19]). The rationale of belief extrapolation is that as long as nothing tells the contrary, fluents do not change. So, our aim is to find the best trajectories satisfying the observations, i.e., the most static trajectories. This is why we introduce preference relations on trajectories:

Definition 4 (preference relations on trajectories)
A preference relation $\preceq$ *is a reflexive and transitive relation on* $TRAJ_N$ *(not necessarily connected). We note*

- $\tau \prec \tau'$ *(read "*τ *is strictly preferred to* τ'*") for (*$\tau \preceq \tau'$ *and not* $\tau' \preceq \tau$*),*
- $\tau \sim \tau'$ *for (*$\tau \preceq \tau'$ *and* $\tau' \preceq \tau$*).*

Let $X \subseteq TRAJ_N$, *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* $\tau' \prec \tau$.

Definition 5
• $\preceq$ *is said* inertial *if and only if*

1. *for any two static trajectories of same length* τ *and* τ' *we have* $\tau \sim \tau'$ *and*
2. *for any static* τ *and any non-static* τ' *(of same length) we have* $\tau \prec \tau'$.

• $\preceq$ *is said to be* change-based *if and only if there exists a reflexive and transitive relation* $\preceq^c$ *on* $\mathcal{F} \times \{1, ..., N\}$ *such that for any two trajectories* τ, τ' *we have* $\tau \preceq \tau'$ *if and only if* $Ch(\tau) \preceq^c Ch(\tau')$.

Inertia says that if there exist static trajectories for a given scenario Σ, then the set of preferred trajectories must be the set of all static trajectories for Σ; it means that if it can be consistently assumed that no change occured, then completion consists in jumping to this conclusion.

The fact that a preference relation is change-based induces a simplification, because the relation is then easier to specify. Intuitive preference relations are most of the time change-based (and actually, almost all examples of preference relations given further on are change-based). Example of preference relation which are not change-based are complex relations where the *context* of the change is relevant (see further).

We are now in position to formally define an extrapolation operator.

Definition 6 (extrapolation operators)
Any preference relation $\preceq$ induces an extrapolation operator $\updownarrow_{\preceq}: \mathcal{S} \rightarrow \mathcal{S}$ defined by:

$$Mod(\Sigma\updownarrow_{\preceq}(t)) = \{\tau(t) \mid \tau \in Min(\preceq, Traj(\Sigma))\}$$

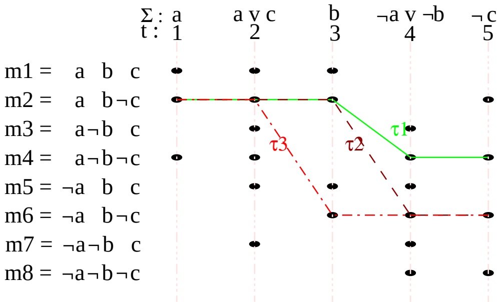
2.3 SOME EXAMPLES OF PREFERENCE RELATIONS

We now give several examples of preference relations. The first one just consists in minimizing the number of elementary changes throughout the process.

- number of changes:

$$\tau \preceq_{nc} \tau' \text{ if and only if } |Ch(\tau)| \leq |Ch(\tau')|$$

Example 1 (continued)



The three trajectories minimal for $\preceq_{nc}$ are:

trajectory τ	$Ch(\tau)$
$\tau_1 = \langle m2, m2, m2, m4, m4\rangle$	$\{\langle\neg b, 4\rangle\}$
$\tau_2 = \langle m2, m2, m2, m6, m6\rangle$	$\{\langle\neg a, 4\rangle\}$
$\tau_3 = \langle m2, m2, m6, m6, m6\rangle$	$\{\langle\neg a, 3\rangle\}$

Extrapolating Σ w.r.t. $\preceq_{nc}$ gives:

$\Sigma\updownarrow_{\preceq_{nc}} = \langle a \wedge b \wedge \neg c,\ a \wedge b \wedge \neg c,\ b \wedge \neg c,\ (a \oplus b) \wedge \neg c,\ (a \oplus b) \wedge \neg c\rangle$;

If the number of changes is minimized fluent by fluent instead of globally, we get the following preference relation:

- number of changes by fluent:

$$\tau \preceq_{ncf} \tau' \text{ iff } \forall f \in \mathcal{F},\ |Ch(\tau, f)| \leq |Ch(\tau', f)|$$

While the latter two preference relations were insensitive to the time points when the fluents changed, the next one is not:

- change set inclusion:

$$\tau \preceq_{csi} \tau' \text{ if and only if } Ch(\tau) \subseteq Ch(\tau')$$

We may also minimize the set of fluents that change at least once:

- inclusion of changing fluents:

$$\tau \preceq_{icf} \tau' \text{ if and only if}$$
$$\{f \in \mathcal{F} | Ch(\tau, f) \neq \emptyset\} \subseteq \{f \in \mathcal{F} | Ch(\tau', f) \neq \emptyset\}$$

A refinement of the global minimzation of the number of changes consists in attaching a penalty to each possible change. The simplest assumption is when this penalty depends only on the identity of the fluent involved in the change. In this case, let $k : \mathcal{F} \rightarrow \mathbb{N}^*$, where $k(f)$ is the penalty induced by a change from $\neg f$ to f, and let $K(\tau) = \sum_t \sum_{\langle f,t\rangle \in Ch(\tau)} k(f)$.Note that we do not necessarily have $k(f) = k(\neg f)$ since some fluents can change more easily than other, for instance $k(alive)$ should be less than $k(\neg alive)$[2].

- penalty:

$$\tau \preceq^k_{cp} \tau' \text{ if and only if } K(\tau) \leq K(\tau')$$

Note that we recover $\preceq_{nc}$ by letting $k(f) = 1$ for all f.

A completely different type of minimization, which has been proposed in many places before, consists in delaying change as much as possible, e.g., preferring trajectories where change occur as late as possible. This policy is called *chronological minimization* in [20]; it is extensively discussed in [19].

- chronological : the preference relation $\preceq_{achr}$ is defined recursively by

[2]As proposed by a reviewer, it could appear more natural to allow no-cost changes, i.e., defining k as a function from $\mathcal{F}$ to $\mathbb{N}$ (instead of $\mathbb{N}^*$). The reason why we impose that changes have a strictly positive penalty is to impose inertia for each fluent. Non-inertial extrapolation would deserve another paper (see some comments in Conclusion).

$$\tau \preceq_{chr} \tau' \text{ if and only if}$$
$$\begin{cases} p = 1 \\ \text{or} \quad Ch(\tau,1) \subset Ch(\tau',1) \\ \text{or} \quad (Ch(\tau,1) = Ch(\tau',1) \\ \qquad \text{and} \quad \langle \tau(2)...\tau(p)\rangle \preceq_{chr} \langle \tau'(2)..\tau'(p)\rangle) \end{cases}$$

Alternatively, one may wish to prefer earlier changes:

- anti-chronological :

$$\tau \preceq_{achr} \tau' \text{ if and only if}$$
$$\begin{cases} p = 1 \\ \text{or} \quad Ch(\tau,p) \subset Ch(\tau',p) \\ \text{or} \quad (Ch(\tau,p) = Ch(\tau',p) \\ \qquad \text{and} \quad \langle \tau(1)..\tau(p-1)\rangle \preceq_{achr} \langle \tau'(1)...\tau'(p-1)\rangle) \end{cases}$$

Lastly, we may want to compare *sequences of changes* independently of their precise location in time: only the changes involved and the order in which they occured are relevant.

- change sequence dominance:
$\tau \preceq_{cseq} \tau'$ if and only if there is an injective mapping F from $Ch(\tau)$ to $Ch(\tau')$ satisfying

1. fluent preservation: $\langle f', t'\rangle = F(\langle f, t\rangle)$ implies $f' = f$;

2. monotonicity w.r.t. time: for any two changes $\langle f_1, t_1\rangle, \langle f_2, t_2\rangle \in Ch(\tau)$, if $\langle f_1', t_1'\rangle = F(\langle f_1, t_1\rangle)$ and $\langle f_2', t_2'\rangle = F(\langle f_2, t_2\rangle)$ then $t_1 < t_2$ if and only if $t_1' < t_2'$.

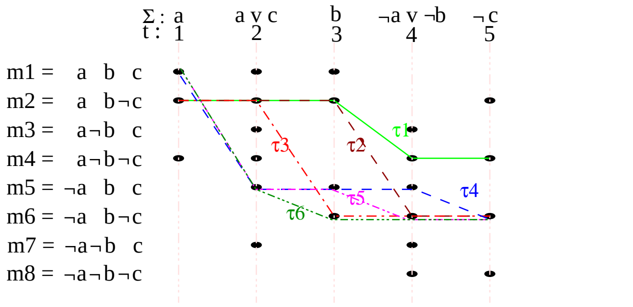
Note that $\tau \sim_{cseq} \tau'$ if and only if the fluents which change in τ and in τ' are the same *and strictly in the same order.*

Example 1 (continued)
The three trajectories minimal for $\preceq_{icf}$ *are the same than for* $\preceq_{nc}$*:* τ_1*,* τ_2 *and* τ_3*.*

There are three more minimal trajectories for $\preceq_{csi}$*, namely:*

trajectory τ	$Ch(\tau)$
$\tau_4 = \langle m1, m5, m5, m5, m6\rangle$	$\{\langle\neg a, 2\rangle, \langle\neg c, 5\rangle\}$
$\tau_5 = \langle m1, m5, m5, m6, m6\rangle$	$\{\langle\neg a, 2\rangle, \langle\neg c, 4\rangle\}$
$\tau_6 = \langle m1, m5, m6, m6, m6\rangle$	$\{\langle\neg a, 2\rangle, \langle\neg c, 3\rangle\}$

As to $\preceq_{cseq}$*, the minimal trajectories are* τ_1*,* τ_2 *and* τ_3*. Notice that* $\tau_2 \prec_{cseq} \tau_4$ *and that* $\tau_4 \sim_{cseq} \tau_5 \sim_{cseq} \tau_6$*. Remark also that* $\tau_7 = \langle m1, m1, m6, m6, m6\rangle$ *cannot be compared w.r.t.* $\sim_{cseq}$ *to* τ_5 *and* τ_6*: indeed,* $Ch(\tau_7) = \{\langle\neg a, 3\rangle, \langle\neg c, 3\rangle\}$*, therefore, in* τ_7 *the changes from* $\neg a$ *to* a *and from* c *to* $\neg c$ *occur simultaneously, whereas for* τ_4*,* τ_5 *and* τ_6*, the changes from* $\neg a$ *to* a *occurs before the change from* c *to* $\neg c$*.*



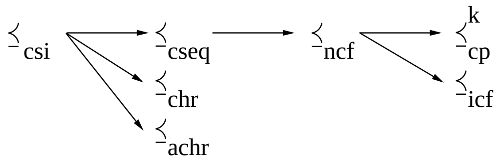
Abbreviating $\updownarrow_{\preceq_x}$ *by* $\updownarrow_x$ *we have:*
$\Sigma\updownarrow_{nc} = \Sigma\updownarrow_{ncf} = \Sigma\updownarrow_{cseq} = \langle a \wedge b \wedge \neg c,\ a \wedge b \wedge \neg c,\ b \wedge \neg c,\ (a \oplus b) \wedge \neg c,\ (a \oplus b) \wedge \neg c\rangle$;

$\Sigma\updownarrow_{csi} = \langle a \wedge b,\ (a \oplus c) \wedge b,\ b \wedge (\neg a \vee \neg c),\ (\neg a \wedge b) \vee (a \wedge \neg b \wedge \neg c),\ (a \oplus b) \wedge \neg c\rangle$;

$\Sigma\updownarrow_{chr} = \langle a \wedge b \wedge c,\ a \wedge b \wedge c,\ a \wedge b \wedge c,\ (a \oplus b) \wedge \neg c,\ (a \oplus b) \wedge \neg c\rangle$;

$\Sigma\updownarrow_{achr} = \langle a \wedge b,\ (a \oplus c) \wedge b,\ \neg a \wedge b \wedge \neg c,\ \neg a \wedge b \wedge \neg c,\ \neg a \wedge b \wedge \neg c\rangle$.

Note that we have the following inclusions between preference relations, represented by this Hasse diagram (edges represent inclusion).



All these preference relations 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* as in [4]. A (deterministic) *event* ε is a mapping from $\mathcal{M}$ to $\mathcal{M}$, where $\varepsilon(m)$ is the result of event ε when the initial state is m. The *empty event* $\varepsilon_\emptyset$ is the identity. Each event has a penalty $k(\varepsilon) \geq 0$ with $k(\varepsilon) = 0$ if and only if $\varepsilon = \varepsilon_\emptyset$.

Given a set of events $\mathcal{E}$ and two models m, m', let $k^{\mathcal{E}}(m, m') = \min\{k(\varepsilon) \mid \varepsilon \in \mathcal{E}, m' = \varepsilon(m)\}$ ($= +\infty$ if no such sequence exists) and $K^{\mathcal{E}}(\tau) = \sum_{t=0..N-1} k^{\mathcal{E}}(\tau(t), \tau(t+1))$. We now define the event-based preference relation:

event penalty

$$\tau \preceq_e^{\mathcal{E}} \tau' \text{ if and only if } K^{\mathcal{E}}(\tau) \leq K^{\mathcal{E}}(\tau')$$

This definition can be refined by assuming that several consecutive event occurences are possible between two time points. This leads us to define an *event sequence* as a combination of events $\theta = \varepsilon_1 \circ ... \circ \varepsilon_p$, such that,

$\theta(m) = \varepsilon_1(\varepsilon_2(...(\varepsilon_p(m))))$; we let $k(\theta) = \sum_{i \in 1..p} k(\varepsilon_i)$. Given a set of events $\mathcal{E}$ and two models m, m', we now let $k^{\mathcal{E}}(m, m') = \min_{\theta | \theta(m) = m'} k(\theta)$ ($= +\infty$ if no such sequence exists). $K^{\mathcal{E}}(\tau)$ and $\preceq_e^{\mathcal{E}}$ are then defined as above. The events (or event sequences) minimizing $k^{\mathcal{E}}(m, m')$ can be considered as the plausible explanations of the transition from m to m'.

Example 1 (continued) *Let $\mathcal{E} = \{\varepsilon_\emptyset, \varepsilon_1, eps_2\}$ where ε_1 switches simulatenously the truth values of a and b: $\varepsilon_1(a, b, c) = (\neg a, \neg b, c)$; $\varepsilon_1(a, b, \neg c) = (\neg a, \neg b, \neg c)$; $\varepsilon_1(a, \neg b, c) = (\neg a, b, \neg c)$ etc. and ε_2 switches the truth value of c. Assume that $k(\varepsilon_1) = k(\varepsilon_2) = 1$. In both cases where we allow for single events only or event sequences, the only preferred trajectory is $\langle a \wedge b \wedge \neg c, a \wedge b \wedge \neg c, a \wedge b \wedge \neg c, \neg a \wedge \neg b \wedge \neg c, \neg a \wedge \neg b \wedge \neg c \rangle$, corresponding to the occurrence of event ε_1 between $t = 3$ and $t = 4$.*

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

2.4 PROPERTIES

We start by giving a set of basic properties and then we show that under weak conditions, they are satisfied by extrapolation operators.

Let Σ be a consistent scenario and φ, ψ two propositional formulas.

E1 $\Sigma\updownarrow \models \Sigma$;
E2 $\Sigma\updownarrow$ is consistent;
E3 if $\Sigma \equiv \Sigma'$ then $\Sigma\updownarrow \equiv \Sigma'\updownarrow$;
E4 $\Sigma\updownarrow(t) \wedge \varphi \models (\Sigma \wedge [t]\varphi)\updownarrow(t)$;
E5 if $\bigwedge_{t' \in 1..N} \Sigma(t')$ is consistent then for all $t \in 1 \dots N$, we have $\Sigma\updownarrow(t) = \bigwedge_{t' \in 1..N} \Sigma(t')$;
E6 if $\Sigma\updownarrow(t) \wedge \varphi$ is consistent then $(\Sigma \wedge [t]\varphi)\updownarrow(t) \models \Sigma\updownarrow(t) \wedge \varphi$

Proposition 1

- *Any operator $\updownarrow_\preceq: \mathcal{S} \to \mathcal{S}$ induced by an inertial preference relation $\preceq$ satisfies the properties E1-E5;*
- *Any operator $\updownarrow_\preceq: \mathcal{S} \to \mathcal{S}$ induced by an inertial, total preference relation $\preceq$ satisfies E1-E6.*

E1 says that observations are reliable. E4 says that extrapolation is monotonic when adding a new observation consistent with the extrapolation of the former observations; E4 and E6 together mean that taking account of this new observation consistent with $\Sigma\updownarrow$ before or after extrapolation gives equivalent results. E5 means that if it can be consistently assumed that nothing changed, then this assumption is made, and therefore the extrapolated belief at each time point is the conjunction of all observations. Note that E1, E2 and E4 entail $\Sigma\updownarrow\updownarrow \equiv \Sigma\updownarrow$, *i.e.*, extrapolation is idempotent.

Beyond these rather undebatable properties, there are other intuitive properties that extrapolation operators may be required to satisfy:

E7 if $\varphi \models \psi$ or $\psi \models \varphi$ then $\forall t \neq |\Sigma| + 1, |\Sigma| + 2$,
$(\Sigma.\varphi.\psi.\Sigma')\updownarrow(t) \equiv (\Sigma.(\varphi \wedge \psi).(\varphi \wedge \psi).\Sigma')\updownarrow(t)$;

E8 $\forall t \neq |\Sigma| + 1$,
$$(\Sigma.\top.\Sigma')\updownarrow(t) \equiv \begin{cases} (\Sigma.\Sigma')\updownarrow(t) & \text{if } t \leq |\Sigma| + 1 \\ (\Sigma.\Sigma')\updownarrow(t-1) & \text{if } t > |\Sigma| + 1 \end{cases}$$

E9 if $(\Sigma.\varphi)\updownarrow(|\Sigma| + 1) \equiv (\varphi.\Sigma')\updownarrow(1)$ then $\forall t$,
$$(\Sigma.\varphi.\Sigma')\updownarrow(t) \equiv \begin{cases} (\Sigma.\varphi)\updownarrow(t) & \text{if } t \leq |\Sigma| + 1 \\ (\varphi.\Sigma')\updownarrow(t - |\Sigma|) & \text{if } t > |\Sigma| + 1 \end{cases}$$

Unlike properties E1-E6, which did not explicitely refer to the flow of time, E7-E8 express important properties related to persistence or independence of changes.

E7 says that if an observation is followed or preceded by a stronger information, then the extrapolated beliefs on other time points would have been the same, had we got twice the stronger information instead. Importantly, this property considers persistence not only forward but also backward: if an observation holds at a given time point, then by default it holds afterwards as well as before.

E8 means that an empty observation has no influence on extrapolated beliefs. An extrapolation operator violating E8 is the (weird) operator based on $\preceq$ s.t. $\tau \preceq \tau'$ if and only if $\forall t \geq 2$, $\{f | \langle f, t \rangle \in Ch(\tau)$ or $\langle f, t+1 \rangle \in Ch(\tau)\} \subseteq \{f | \langle f, t \rangle \in Ch(\tau')$ or $\langle f, t+1 \rangle \in Ch(\tau')\}$.

E9 is a kind of Markov property and implies that an extrapolation problem can sometimes be decomposed into independent subproblems; especially, the condition is verified as soon as there is a *complete* observation (having a single model) at $t = |\Sigma| + 1$.

Proposition 2

- *$\updownarrow_{nc}$ and $\updownarrow_{cp}^k$ satisfy E1-E9.*
- *$\updownarrow_{csi}$, $\updownarrow_{cseq}$ and $\updownarrow_{ncf}$ satisfy E1-E5 and E7-E9.*
- *$\updownarrow_{chr}$ and $\updownarrow_{achr}$ satisfy E1-E5, E8 and E9.*
- *$\updownarrow_{icf}$ satisfies E1-E5, E7 and E8.*

- $\updownarrow_e^E$ *satisfies E1-E6, E8 and E9.*

Another interesting property is the *reversibility* of extrapolation. We first need to define the *reversal* of a scenario or a trajectory.

Definition 7
For any scenario $\Sigma = \langle\Sigma(1), .., \Sigma(N)\rangle$, *we define the reversal of* Σ *by*

$$Reverse(\Sigma) = \langle\Sigma(N), .., \Sigma(1)\rangle$$

For any trajectory τ *we define the reversal of* τ *by*

$$Reverse(\tau)(t) = \tau(N - t)$$

Now, we introduce the *reversibility* property:

(R) $Reverse(\Sigma\updownarrow) = (reverse(\Sigma))\updownarrow$.

Proposition 3 $\updownarrow_{\preceq}$ *satisfies (R) if and only if* $\preceq$ *is such that for all* τ, τ' *of same length* N *we have* $\tau \preceq \tau'$ *if and only if* $Reverse(\tau) \preceq Reverse(\tau')$

Proposition 4

- $\updownarrow_{csi}$, $\updownarrow_{ncf}$, $\updownarrow_{nc}$ *and* $\updownarrow_{icf}$ *satisfy (R);*
- $\updownarrow_{cp}^k$ *satisfies (R) if and only if* $\forall f \in \mathcal{F}$ *we have* $k(f) = k(\neg f)$*;*
- $\updownarrow_{chr}$ *and* $\updownarrow_{achr}$ *do not satisfy (R);*
- $\updownarrow_{ep}^E$ *satisfies (R) if and only if* $\forall m, m'$ *we have* $k^{\mathcal{E}}(m, m') = k^{\mathcal{E}}(m', m)$

3 COMPUTATIONAL ISSUES

3.1 COMPLEXITY

We start by identifying the complexity of belief extrapolation for each of the preference relations given above. We assume that the reader is familiar with the complexity classes Δ_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 e.g. [18]).

Proposition 5 EXTRA$_x$ *is the decision problem: given* Σ, ψ *and* t, *decide whether* $\Sigma \updownarrow_x (t) \models \psi$ *holds. Then we have the following:* EXTRA$_{csi}$, EXTRA$_{icf}$, EXTRA$_{cseq}$, EXTRA$_{chr}$ *and* EXTRA$_{achr}$ *are* Π_2^p*-complete;* EXTRA$_{cp}^k$ *is in* Δ_2^{P} [3]*;* EXTRA$_{nc}$ *is in* $\Delta_2^{\mathsf{P}}[\mathcal{O}(\log n)]$.

[3] This result has to be drawn together with the complexity of inference under the backwards semantics in [17] since both problems are closely related (see Section 6).

3.2 COMPUTING EXTRAPOLATED BELIEFS

For the rest of this Section we fix the preference relation to $\preceq_{csi}$.

We define now a *minimal explanation* γ for Σ as the change set $Ch(\tau)$ of a $\preceq_{csi}$-preferred trajectory τ of Σ.

The information provided by the set of all minimal explanations is very rich: even if it does not always determine the preferred trajectories in the general case (because there may be several trajectories having the same change set), it does as soon as there is a t_i such that $\Sigma(t_i)$ is *complete* (i.e., has only one model m_i), and in this case, from the set of minimal explanations together with m_i we can decide in polynomial time whether $\Sigma\updownarrow(t) \models \psi$ holds for any t and ψ.

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 2 $\Sigma = \langle a \wedge b \wedge c \wedge d,\ \top,\ \neg b,\ d,\ \neg a,\ \top,\ b,\ \neg c \vee \neg d\rangle$ *has exactly* $4 \times 2 \times 4 \times (7+4) = 352$ *minimal explanations (and as many minimal trajectories). Now, these 352 minimal explanations can be expressed compactly:* γ *is a minimal explanation for* Σ *if and only if it contains (i)* $\langle\neg a, t\rangle$ *for (exactly) one* $t \in \{2..5\}$ *(meaning that a became false exactly once, at 2, 3, 4 or 5 – since it was true at 1 and false at 5), (ii)* $\langle\neg b, t\rangle$ *for one* $t \in \{2, 3\}$, $\langle b, t\rangle$ *for one* $t \in \{4..7\}$, *and either (iii)* $\langle\neg c, t\rangle$ *for one* $t \in \{2..8\}$ *or (iii')* $\langle\neg d, t\rangle$ *for one* $t \in \{5..8\}$ *(and nothing else).*

To capture this simple notion we introduce the following notion: a *compact change set* Γ is a set $\{\langle f_i, t_i^-, t_i^+\rangle \mid i = 1..p\}$ where the f_i are fluents and the (t_i^-, t_i^+) are pairs of time points such that $0 < t_i^- \leq t_i^+ \leq N$. The set of change sets covered with Γ is $Cov(\Gamma) = \{\langle f_1, t_1\rangle, ..., \langle f_p, t_p\rangle \mid \forall i = 1..p, t_i^- \leq t_i \leq t_i^+\}$. We say that Γ *explains* Σ if and only if every $\gamma \in Cov(\Gamma)$ explains Σ.

Intuitively, a compact change set consists of the disjunction of all the change sets it covers. We say that compact change set Γ *explains minimally* Σ if and only if $\forall\gamma \in Cov(\Gamma)$, γ is a minimal explanation for Σ.

Γ is a *maximally compact 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 $Cov(\Gamma') \subset Cov(\Gamma)$. The set of MCMEs for Σ expresses minimal explanations in a compact way and is redundancy-free: indeed, it can be easily shown that for each minimal explanation γ for Σ there is one and only one MCME Γ for Σ covering γ. In the previous

example, there are two MCMEs:
$\{\langle\neg a,2,5\rangle,\langle\neg b,2,3\rangle,\langle b,4,7\rangle,\langle\neg c,2,8\rangle\}$ and
$\{\langle\neg a,2,5\rangle,\langle\neg b,2,3\rangle,\langle b,4,7\rangle,\langle\neg d,5,8\rangle\}$.

When observations are disjunction-free we easily get:

Proposition 6 *If for each t, $\Sigma(t)$ is a consistent conjunction of fluents, then there is only one maximally compact minimal explanation for Σ.*

Compact minimal explanations can be computed by adding to Σ a maximal set of persistence assumptions expressing that each fluent f persists by default during intervals between two consecutive observations "relevant" to f. This notion of relevance is modeled in this simple way (see [14]): We say that an observation φ is *relevant* to a propositional variable v, denoted $Rel(\varphi, f)$, if and only if there is no formula equivalent to φ in which v does not appear at all[4].

Now, for a variable v, let $R_\Sigma(v) = \{t \in 1..N \text{ s.t. } Rel(\Sigma(t), v)\}$; now, the set of *relevant intervals* $RI_\Sigma(v)$ to v w.r.t. Σ is defined by $RI_\Sigma(v) = \{[t,t'] \mid t \in R_\Sigma(v), t' \in R_\Sigma(v), [t+1,t'-1] \cap R_\Sigma(v) = \emptyset\}$. The intuitive meaning of relevant intervals is that it is sound to consider only persistence assumptions regarding a given variable within these intervals only[5].

Variables are then time-stamped by the time point they refer to[6]: if $v \in \mathcal{V}$ and $t \in 1..N$ then v_t holds if and only if v holds at time t. If ϕ is a formula built on $\mathcal{V}$ then ϕ_t is obtained from ϕ by replacing each variable v in ϕ by v_t. Let $\Sigma^* = \{\Sigma(t)_t \mid t = 1..N\}$ and let the set of *persistence assumptions* for Σ be $P_\Sigma = \bigcup_{\substack{v \in \mathcal{V} \\ [t,t'] \in RI_\Sigma(v)}} \{v_t \rightarrow v_{t'}, \neg v_t \rightarrow \neg v_{t'}\}$. Define the *P-explanations* for $\langle P_\Sigma, \Sigma^*\rangle$ as the minimal subsets π of P_Σ such that $P_\Sigma \setminus \pi$ is consistent with Σ^*. Each P-explanation π can be identified with a compact change set $\Gamma(\pi)$ in an obvious way; for instance, $\pi = \{a_4 \rightarrow a_7, \neg c_5 \rightarrow \neg c_6\}$ draws $\Gamma(\pi) = \{\langle\neg a,5,7\rangle,\langle c,6,6\rangle\}$, i.e., a became false once between 4 and 7 (i.e., at 5, 6, or 7) and c became true at 6. We have the following:

Proposition 7

- *for each P-explanation π for $\langle P_\Sigma, \Sigma^*\rangle$, each change set $\gamma \in \Gamma(\pi)$ is a minimal explanation for Σ* [7]*;*
- *each minimal explanation for Σ is covered by exactly one P-explanation, i.e., there exists one and only one π such that $\gamma \in \Gamma(\pi)$.*

Thus, a covering set of compact explanations can be computed using abduction algorithms. We also point that complexity results and tractable classes about logic-based abduction [9], can be used to evaluate the difficulty of this task and find some tractable subclasses of extrapolation.

4 EXTRAPOLATION AND REVISION

We do not recall here the well-known AGM postulates [1], which are anyway not crucial to the understanding of this Section.

Belief revision usually deals with static worlds, and therefore works with formulas pertaining to the same time point. However, as remarked in [11], "what is important for revision is not that the world is static, but that the propositions used to describe the world are static". This explains why, when indexing propositional variables by time points as we do in the Section 3, we manage to capture a fully dynamic process within a pure belief revision framework: belief extrapolation is equivalent to revising the prior belief that all fluents persist throughout time by the observations.

4.1 FROM EXTRAPOLATION TO REVISION

A first connection appears when we consider sequences of *two* observations only. Extrapolation operators on sequences of length 2 are called *2-extrapolation* operators.

Since the language is finite, we consider revision operators as acting on propositional formulas as in [13] (and not on belief sets). From $\updownarrow: \mathcal{L}^2 \rightarrow \mathcal{L}^2$ we induce the belief change operator $\star = \Phi(\updownarrow) : \mathcal{L}^2 \rightarrow \mathcal{L}^2$ by $KB \star \alpha = \langle KB, \alpha\rangle \updownarrow (2)$. Conversely, from a belief change operator $\star$ we induce the operator $\updownarrow = \Psi(\star)$ by $\langle\Sigma(1),\Sigma(2)\rangle \updownarrow (2) = \Sigma(1) \star \Sigma(2)$ and $\langle\Sigma(1),\Sigma(2)\rangle \updownarrow (1) = \Sigma(2) \star \Sigma(1)$. Note that, obviously, $\Phi(\Psi(\star)) = \star$, but $\Psi(\Phi(\updownarrow) = \updownarrow$ does not necessarily hold.

[4] or equivalently, φ is *not* relevant to v if and only if $\varphi_{v \leftarrow \top}$, $\varphi_{v \leftarrow \bot}$ and φ are logically equivalent.

[5] This method could be further improved by distinguishing between *positive* and *negative* observations.

[6] A similar translation in used in [11] so as to represent beliefs about different time point in a pure belief revision framework.

[7] 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 $\{\langle a,4,6\rangle\}$ is a MCME. If we wish to obtain the MCMEs for Σ we have to define a compactification procedure whose precise description is outside the scope of the paper.

Proposition 8

1. $\updownarrow$: $\mathcal{L}^2 \to \mathcal{L}^2$ *satisfies E1-E6 if and only if* $\Phi(\updownarrow)$ *is an AGM belief revision operator;*
2. $\star$ *is an AGM belief revision operator if and only if* $\Psi(\star)$ *is a 2-extrapolation operator satisfying E1-E6.*

This result sheds more light on what belief extrapolation intuitively consists of. Assuming $\preceq$ c! connected and $N = 2$, the extrapolated belief set at time 2 consists of the revision of beliefs at time 1, which we believe to persist to time 2, by the observations at time 2 which hold for certain. This looks like a pure revision framework, but have in mind that the world is dynamic. Symmetrically, the extrapolated belief set at time 1 consists of the revision of beliefs at time 2, which we believe to " persist backwards" to time 1, by the observations at time 1.

4.2 FROM REVISION TO EXTRAPOLATION

Proposition 9 *Let $\star$ be an AGM belief revision operator defined on a propositional language generated by $\mathcal{V} = \{v_1, .., v_n\}$. We index variables by time points as in Section 3. Let $PERS = \bigwedge_{\substack{v \in \mathcal{V} \\ t=1..N-1}} v_t \leftrightarrow v_{t+1}$ and $\Sigma^* = \{\Sigma(t)_t \ |t = 1..N\}$. Then the operator $E : \Sigma \mapsto E(\Sigma)(t) = \psi_t$ where ψ_t is the strongest formula mentioning only variables indexed by t such that $PERS \star \Sigma^* \models \psi_t$, is an extrapolation operator satisfying E1-E6.*

Thus, we can draw an extrapolation operator from any revision operator. Actually, it can be shown that properties E1-E6 correspond to a reformulation of the AGM postulates so as to take account from different time points; however, they are only little relevant to the specificities concerned with the flow of time, such as forms of Markovianity and independence of changes expressed by E7-E9. Therefore, as soon as we require other properties than the basic ones, extrapolation is "more specific" than AGM revision.

4.3 EXTRAPOLATION AND ITERATED REVISION

Now, an interesting question is the possible connections between extrapolation and *iterated belief revision*. Especially, Lehmann's iterated revision [16] considers sequences of consistent formulas[8], therefore the input is exactly as ours. Since he 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 $[\Sigma]$, with the projection of the extrapolation at the final time point $\Sigma \updownarrow (N)$ (where $N = |\Sigma|$).

Proposition 10 *Let $\updownarrow$ be a basic extrapolation operator. The operator $\Sigma \mapsto \Sigma \updownarrow (N)$ satisfies postulates I1, I2, I3, I5 and I6 of [16].*

On the other hand, I4 is generally not satisfied[9] and I7 is never satisfied. 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" [16]. The difference between extrapolation and iterated revision is clear when considering the following example: let $\Sigma = \langle a \to b, a, \neg a \rangle$; any "reasonable" extrapolation operator satisfies $\Sigma\updownarrow = \langle a \wedge b, a \wedge b, \neg a \wedge b \rangle$ (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.*, [7] [2]) give $[a \to b, a, \neg a] = \neg a$. The reason for this difference is that iterated revision is concerned with pieces of information concerning a *static world*, 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).

5 EXTRAPOLATION AND UPDATE

An update operator is a function which associates to 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. As for revision, Katsuno and Mendelzon have proposed a set of postulates that an update operator $\diamond$ should satisfy. The most "typical" postulate is

U8 $(K1 \vee K2) \diamond \varphi \leftrightarrow (K1 \diamond \varphi) \vee (K2 \diamond \varphi)$.

which tells that models of K are updated indepen-

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

[9] It is satisfied if and only if $\preceq$ satisfies the property: for any $\Sigma = \langle \varphi_1, ...\varphi_N \rangle$ and any $\tau = \langle \tau(1)..\tau(N) \rangle$, $\tau \in Min(\preceq, Traj(\Sigma))$ implies $\langle \tau(1)..\tau(N-1) \rangle \in Min(\preceq, Traj(\langle \varphi_1, ..., \varphi_{N-1} \rangle))$. Only chronological operators such as $\updownarrow_{chr}$ satisfy this property.

dently.

5.1 EXTRAPOLATION IS NOT BELIEF UPDATE

When $N = 2$, belief extrapolation projected 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. This may look similar to belief update, however 2-extrapolation *is not* belief update. This is clear from point 1 of Proposition 8 (or from E5): it is known that the AGM postulates are inconsistent with postulate U8 [12] as soon as the language contains at least two propositional symbols[10]:

Proposition 11 *Assume that the language contains at least two propositional symbols. Let $N = 2$, $\updownarrow$ a 2-extrapolation operator, and $\star$ the induced belief change operator as in Section 4. Then $\star$ cannot satisfy U8. Therefore, no belief change operator induced by a 2-extrapolation operator is a belief update operator.*

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. This inadequacy of update to handle observations is not new (see e.g. [4]), but is rarely mentioned in papers about belief update: indeed, saying that updating KB by α comes down to incorporate observation α (coming from sensors for instance) reflecting a change in the world is not correct.

So, the impossibility (with update) to infer new beliefs at time 1 from the incorporated belief at time 2 means that we do not *observe* (from sensors, by communicating with other agents etc.) anything at time 2[11]. Therefore, if not an observation, the input α is necessarily a *prediction*, more precisely, the projection of the expected effects of some "explicit change", most likely, the *expected* (not the *observed*) effect of an action.

[10] More precisely, U8 cannot be satisfied together with the strong K$\star$3 and K$\star$4 in presence of the more harmless postulates K$\star$2, K$\star$5 and K$\star$6.

[11] However, a second interpretation would be acceptable: there are observations at time 2 but they are completely unrelated to the world at time 1 (which implies giving up persistence assumptions and minimal change) and a belief change operator in this extreme case is not worth studying.

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 observation. In Sandewall's taxonomy [19], extrapolation is adequate for the action-free subclass of K-**IS** (correct knowledge, inertia and surprises) while update is adequate for the class $\mathrm{K_p}$-**IA** (no observations after the initial time, inertia and alternative results of actions).

5.2 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 rather easily as follows. A *mixed scenario* now contains both observations and explicit updates (corresponding to expected effects of actions) and is therefore a sequence of formulas labelled either by `observe` or by `update`, the first member of the sequence being of the `observe` type. For instance, $\Sigma_0 = \langle$ `observe`: $a \vee b$, `update`: $\neg b$, `observe`: $\neg a \rangle$ is a mixed scenario expressing that $a \vee b$ was first observed, then an action was performed whose expected effect is $\neg b$, and, after it, $\neg a$ was observed.

Let $\diamond$ be an arbitrary update operator and $\preceq$ a change-based preference relation induced by a relation $\preceq^c$ on change sets. Let Σ be a mixed scenario of length N and $\tau = \langle M_0, ..., M_N \rangle$ a trajectory. We say that τ satisfies Σ with respect to $\diamond$ if and only if (1) for every t such that $\Sigma(t) = \mathtt{observe}{:}\alpha$ we have $M_t \models \alpha$; (2) for every t such that $\Sigma(t) = \mathtt{update}{:}\alpha$ we have $M_t \in M_{t-1} \diamond \alpha$. Let $Traj_\diamond(\Sigma)$ the set of trajectories satisfying Σ w.r.t $\diamond$. Let $UnprCh(\tau)$ the set of unpredicted changes of τ, i.e., the changes $\langle f, t \rangle \in Ch(\tau)$ such that f is not involved in an update at time t. τ is a *preferred trajectory* with respect to $\diamond$ and $\preceq$ if and only if there is no trajectory $\tau' \in Traj_\diamond(\Sigma)$ such that $UnprCh(\tau') \prec^c UnprCh(\tau')$. Lastly, the operator $\Updownarrow$ induced by $\preceq$ and $\diamond$ is defined by $Mod(\Sigma \Updownarrow (t)) = \{\tau(t) \mid \tau \in Pref(Traj_\diamond(\Sigma))\}$. Therefore, preferred trajectories are selected in two steps: (1) filter out trajectories that do not satisfy all observations or all explicit changes; (2) among remaining trajectories, minimize unexpected change. Choosing any $\diamond$ satisfying the KM postulates and any $\preceq$, the latter Σ_0 yields $\Sigma_0 \Updownarrow = \langle a \vee b, \neg b, \neg a \wedge \neg b \rangle$.

This integration of extrapolation and update can deal with observations and actions (deterministic or not). Three issues are not dealt with:

1. *unreliable observations*, or misperceptions. These could be handled by attaching entrenchment levels to observations and then apply a temporal generalization of knowledge transmutations [21]. See also [5] and [6] for alternative ways of handling unreliable observations.

2. *normal and exceptional effects of actions*;

3. *complex events, whose expected effects yield an update, but whose occurrence is not known a priori but (abductively) inferred from the observations.* This issue is handled (at least for sequences of length 2) by Boutilier's *generalized update* [4].

All three issues could be dealt with by attaching degrees of surprise to misperceptions, event occurrences, action effects and event effects, and and then maximize the global level of surprise over the whole trajectory. This is a topic for further research.

6 RELATED WORK

In addition to the position of extrapolation with respect to revision, iterated revision, update and generalized update, already discussed in Sections 4 and 5, several works are related to belief extrapolation.

6.1 BERGER, LEHMANN AND SCHLECHTA

Berger et al. [3] propose a generic class of belief change operators which is actually a subclass of the set of extrapolation operators. More exactly, an extrapolation operator is a belief change operator as in [3] if and only if its preference relation satisfies the following property (BLS). Let $h(\tau)$ defined from τ 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 so on until such an integer can no longer be found. The property (BLS) is then the following: for any trajectories τ, τ' such that there exists a sequence of integers $1 \leq j_1 < j_2 < ... < j_k$ such that $h(\tau) = \langle m_0, ..., m_n \rangle$ and $h(\tau') = \langle m_{j_0}, ..., m_{j_k} \rangle$, we have $\tau \preceq \tau'$. It can be shown that, among the preference relations we proposed, those that satifsy (BLS) are $\preceq_{nc}$, $\preceq_{ncf}$, $\preceq_{ifc}$, $\preceq_{cseq}$ and $\preceq^k_{cp}$. While they do not study various preference relations on trajectories nor investigate computational issues, Berger et al. get a both-directions representation result; we have not got such a result yet for belief extrapolation (note that our operators being less constrained than theirs, the latter representation result does not apply to belief extrapolation).

6.2 LIBERATORE AND SCHAERF

Liberatore and Schaerf [17] 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 $\preceq_{cp}$. The other semantics ("pointwise") yields iterated update (but is incompatible with extrapolation because of U8 which underlies this semantics); it seems that the authors did not consider mixing updates and observations by using both semantics simultanesouly as we briefly did in Section 5. On the other hand, BReLS provides an integrated treatment of both static and dynamic reasoning, and thus includes the possibility of reprsenting exceptional effects or misperceptions.

6.3 OTHER WORKS

Friedman and Halpern [11] define a very general framework for belief change, based on trajectories (or "runs") and in which the central process is that of *conditioning*. Like belief revision and update, it can be shown that extrapolation is an instance of their general framework.

Largouët et Cordier [15] propose a (probabilistic) framework for belief change integrating observations and expectations allowing, like ours, to complete an initial scenario by persistence assumptions. Their completion process operates in two subsequent steps: beliefs are first propagated forwards (from time 0 to time n) and then propagated backwards (from time N to time 0). Their process, by performing postdiction steps *after* prediction steps, give priority to backwards persistence and therefore almost never gives the same results as belief extrapolation.

In an earlier work [8] we considered a probabilistic model of unpredicted change which may be seen as the probabilistic counterpart of the approach given here. A prior probability is attached to each elementary change and assuming prior independence between occurrences of elementary changes. The case where probabilities of change are infinitely small was particularly focused on.

7 CONCLUSION

We have proposed 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 (and located both families of operators within Sandewall's taxonomy).

Further directions of work include :

- applying belief extrapolation to temporal diagnosis;
- developing a full framework for reasoning about both predicted change (action effects) and unpredicted change;
- generalize extrapolation to partially non-inertial systems where some fluents have a specific dynamics.

Acknowledgements

We would like to thank two anonymous reviewers for their useful comments, as well as Marie-Odile Cordier, Andreas Herzig, Sébastien Konieczny, Pierre Marquis and Odile Papini for helpful discussions.

References

[1] C. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change : partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.

[2] S. Benferhat, S. Konieczny, O. Papini, and R. Pino-Pérez. Iterated revision by epistemic states: axioms, semantics and syntax. In *Proceedings of ECAI-2000*, pages 13–17, 2000.

[3] S. Berger, D. Lehmann, and K. Schlechta. Preferred history semantics for iterated updates. *Journal of Logic and Computatation*, 9(6):817–833, 1999.

[4] C. Boutilier. A unified model of qualitative belief change: a dynamical systems perspective. *Artificial Intelligence*, 1-2:281–316, 1998.

[5] C. Boutilier, N. Friedman, and J. Halpern. Belief revision with unreliable observations. In *Proceedings of AAAI-98*, pages 127–134, 1998.

[6] S. Coradeschi. Reasoning with misperception in the features and fluents framework. In *Proceedings of ECAI-96*, pages 657–661, 1996.

[7] A. Darwiche and J. Pearl. On the logic of iterated revision. *Artificial Intelligence*, 89:1–29, 1997.

[8] F. Dupin de Saint-Cyr and J. Lang. Reasoning about unpredicted change and explicit time. In *Lecture Notes in Artificial Intelligence (Proc. of ECSQARU-97)*, 1997.

[9] Th. Eiter and G. Gottlob. The complexity of logic-based abduction. *Journal of the Association for Computing Machinery*, 42(1):3–42, 1995.

[10] N. Friedman and J.Y. Halpern. Belief revision: A critique. In *Proc. of the 6^{th} KR*, pages 421–431, 1996.

[11] N. Friedman and J.Y. Halpern. Modelling beliefs in dynamic systems. part ii: revision and update. *JAIR*, 10:117–167, 1999.

[12] H. Katsuno and A.O. Mendelzon. On the difference between updating a knowledge base and revising it. In *Proc. of KR'91*, pages 387–394, 1991.

[13] H. Katsuno and A.O. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52:263–294, 1991.

[14] J. Lang and P. Marquis. Complexity results for independence and definability in propositional logic. In *Proc. of KR'98*, pages 356–367, 1998.

[15] C. Largouët and M.O. Cordier. Combining observations and expectations: application to the refinement of an image sequence classification. In *UAI workshop on Fusion of domain knowledge with data for decision support*, Standford, CA, USA, 2000.

[16] D. Lehmann. Belief revision, revised. In *Proc. of IJCAI'95*, pages 1534–1540, 1995.

[17] P. Liberatore and M. Schaerf. Brels: A system for the integration of knowledge bases. In *Proc. of KR'00*, pages 145–152, 2000.

[18] C. H. Papadimitriou. *Computational Complexity.* Addison-Wesley, 1994.

[19] E. Sandewall. *Features and Fluents.* Oxford University Press, 1995.

[20] Y. Shoham. *Reasoning about Change - Time and Causation from the Standpoint of Artificial Intelligence.* MIT Press, 1988.

[21] Mary-Anne Williams. Transmutations of knowledge systems. In *Proceedings of KR'94*, pages 619–629, 1994.