Extrapolation de croyances (ou comment raisonner avec des observations et des changements imprévus)

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

F. Dupin de Saint-Cyr[1] J. Lang[2]

[1] LERIA, Université d'Angers (France), Florence.Bannay@univ-angers.fr
[2] IRIT, Université Paul Sabatier, Toulouse (France), lang@irit.fr

Résumé

Nous donnons un nouveau cadre logique pour raisonner avec des observations à différents instants. On appelle "extrapolation" le processus cohérent à compléter les croyances initiales, provenant d'observations, par une hypothèse de changement minimal. Nous donnons une sémantique générale et proposons plusieurs opérateurs d'extrapolation. Nous étudions quelques propriétés vérifiées par ces opérateurs, identifions leur complexité, et donnons quelques éléments algorithmiques. Ensuite, nous étudions en détail la position de l'extrapolation de croyances par rapport à la révision et à la mise-à-jour; en particulier, nous montrons que, bien qu'elle concerne un monde non-statique, l'extrapolation de croyances a plus à voir avec la révision qu'avec la mise-à-jour, et nous montrons comment intégrer mise-à-jour et extrapolation.

Mots Clef

Raisonnement non monotone, représentation des connaissances, révision, mise-à-jour.

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, we identify their complexity and give some algorithmic elements. We study in detail the position of belief extrapolation with respect to revision and update ; we show that, although it deals with non-static worlds, belief extrapolation shares more with belief revision than with belief update and we show how to integrate update and extrapolation.

Keywords

Nonmonotonic reasoning, knowledge representation, revision, update.

1 Introduction

La distinction entre révision et mise-à-jour [11] est généralement exprimée en mettant en avant l'opposition entre monde statique et monde dynamique : la révision incorpore de nouvelles croyances à d'anciennes, toutes ces croyances portant sur un même monde (statique), tandis que la mise-à-jour modifie les croyances sur l'état du monde *en réponse à un changement du système, "notifié" par une nouvelle formule* φ . La définition informelle que l'on vient de donner de la révision n'est pas ambiguë, par contre, pour la mise-à-jour, il y a une ambiguïté dans la nature exacte de la "nouvelle formule" (en anglais "input formula") φ. Que veut dire exactement une "notification de changement?" On peut penser à plusieurs choses : ainsi, φ peut être une observation faite par l'agent quelque temps après la dernière observation du système; ou encore, φ peut être la conséquence attendue d'une action réalisée depuis la dernière observation du système (ou la conséquence attendue d'un événement dont l'occurrence a été détectée). Si on ne distingue pas formellement ces différents types d'information, on peut obtenir des résultats contre-intuitifs, comme en témoigne cet exemple :

On considère un système à diagnostiquer, dont les composants peuvent tomber en panne, de manière indépendante[1]. Scénario 1*: À l'instant initial $t = 1$, on sait que (i) exactement l'un des deux composants a et b est en panne, et que (ii) le composant c fonctionne correctement : $K = (ab(a) \oplus ab(b)) \wedge \neg ab(c)$*[2]*. On laisse passer un certain temps, et ensuite (à l'instant $t = 2$), de nouvelles observations nous apprennent que b fonctionne bien (on peut ima-*

1. c'est-à-dire qu'l n'y a pas de lien causal a priori entre une panne d'un composant et une panne d'un autre. Pour simplifier le problème, on suppose aussi que le fait qu'un composant en panne se mette à fonctionner normalement n'est pas plus exceptionnel que le fait qu'un composant qui fonctionnait tombe en panne. Cette hypothèse n'est pas si irréaliste si l'on considère que d'autres agents peuvent réparer une panne à l'insu de l'agent dont les croyances nous intéressent.

2. Le prédicat $ab(x)$ signifie que le composant x est anormal.

giner que l'on vient de le tester) mais que c est en panne : $\varphi = \neg ab(b) \wedge ab(c)$.

Scénario 2*: à l'instant* $t = 1$, *comme pour le scénario 1,* $K = (ab(a) \oplus ab(b)) \wedge \neg ab(c)$. *On effectue ensuite l'action de remplacer le composant* b *par un neuf (dont on sait qu'il est en état de marche), ainsi que (malencontreusement) l'action de casser le composant* c. *Les effets cumulés de ces deux actions se traduisent par la formule* $\varphi = \neg ab(b) \wedge ab(c)$.

Ces deux scénarios sont différents et doivent mener à des conclusions différentes. Pour le scénario 1, la façon intuitivement satisfaisante de changer les croyances est la suivante : "puisque le composant b fonctionne correctement, et que les changements sont rares, il y a d'excellentes raisons de croire qu'il fonctionnait déjà correctement à l'instant 1 et qu'à ce même instant, c'était a qui était en panne; comme, à nouveau, les changements sont rares, il y a d'excellentes raisons de croire que a est toujours en panne à l'instant 2. Quant à c que l' on vient d'observer en panne, pas de doute, il l'est." L'état de croyances à $t = 2$ doit donc, en principe, être $ab(a) \wedge \neg ab(b) \wedge ab(c)$. Pour le scénario 2, les choses sont différentes : l'effet du remplacement de b n'apporte aucune information sur son état de fonctionnement antérieur, ni par conséquent sur celui de a. L'état de croyances à $t = 2$ doit donc, en principe, être $\neg ab(b) \wedge ab(c)$ (les croyances concernant a restant inchangées).

La mise-à-jour est le processus correct pour le scénario 2, mais pas pour le scénario 1. Le principal responsable de cet échec de la mise-à-jour à traiter correctement le scénario 1 est la propriété la plus "typique" de la mise-à-jour, à savoir U8 [3], qui dit que les modèles de K doivent être mis-à-jour indépendemment; la plupart des opérateurs de mise-à-jour $\diamond$ satisfont donc $K \diamond \neg ab(b) = \neg ab(b)$, ce qui est le résultat attendu pour le scénario 2 mais pas pour le scénario 1.

À partir de ces deux scénarios on peut avancer une hypothèse : *la mise à jour de croyances n'est pas appropriée quand la nouvelle formule est une observation*. Deux questions en découlent : (Q1) Dans quelles circonstances est-elle appropriée ? (Q2) Comment faut-il changer une base de croyances afin de prendre en compte une observation?

En ce qui concerne Q1, le scénario 2 suggère que la mise à jour de croyances est adaptée lorsque *la nouvelle formule est le résultat attendu d'une action (ou d'un événement) dont l'occurrence est connue de l'agent*. Q2, quant à elle, va mener à la définition d'un nouvel opérateur que nous appelons *extrapolation de croyances*. Cet opérateur prend une séquence d'observations supposées fiables, et les projette vers le passé et vers le futur, en supposant que la valeur de vérité des fluents [4] tend à persister.

En section 2 nous définissons l'extrapolation de croyances, à l'aide d'une sémantique en termes de trajectoires minimales; nous proposons plusieurs opérateurs d'extrapolation et étudions leurs propriétés. En Section 3 nous étudions les propriétés calculatoires de l'extrapolation de croyances : après avoir identifié la complexité des différents opérateurs proposés, nous proposons une méthode de calcul pour le plus simple de ces opérateurs. Ensuite, nous montrons que l'extrapolation, bien qu'elle s'applique à un monde en évolution, est plus proche de la révision que de la mise à jour ; il faut plutôt voir l'extrapolation et la mise à jour comme complémentaires, et un cadre complet pour le raisonnement sur des croyances à travers le temps nécessite une intégration de l'extrapolation et de la mise-à-jour, que nous esquissons. Enfin, nous étudions les liens entre l'extrapolation et plusieurs autres approches.

2 Extrapolation de croyances

2.1 Préliminaires

On considère un langage propositionnel $\mathcal{L}$, construit à partir d'un ensemble fini $\mathcal{V} = \{v_1, ..., v_n\}$ de variables propositionnelles. On note $\mathcal{M} = 2^{\mathcal{V}}$ l'ensemble des interprétations pour $\mathcal{V}$. $Mod(\phi)$ dénote l'ensemble des modèles de la formule ϕ. Un *fluent* f est un littéral, c'est-à-dire une variable ou sa négation : $\mathcal{F} = \{v_1, \neg v_1, ..., v_n, \neg v_n\}$. Si $f = v_i$ (respectivement $\neg v_i$) alors $\neg f$ est égal à $\neg v_i$ (respectivement v_i).

Les résultats de complexité que nous donnons dans cet article font mention de certaines classes de complexité de la hiérarchie polynomiale (voir [19] pour une synthèse). Nous supposons que le lecteur connaît les classes P, NP et coNP. $\Delta_2^p = \mathsf{P}^{\mathsf{NP}}$ est la classe des langages reconnaissables en temps polynomial par une machine de Turing déterministe munie d'oracles NP, où un oracle NP résout une instance quelconque d'un problème dans NP ou coNP en temps unité. $\Theta_2^p = \Delta_2^p[\mathcal{O}(log\ n)]$ est la classe de tous les langages reconnaissables en temps polynomial par une machine de Turing déterministe utilisant un nombre d'appels *borné par une fonction logarithmique* à un oracle NP. Enfin, $\Sigma_2^p = \mathsf{NP}^{\mathsf{NP}}$ est la classe des langages reconnaissables en temps polynomial sur une machine de Turing non-déterministe munie d'oracles NP, et $\Pi_2^p = \mathsf{co}\Sigma_2^p$.

On peut noter que $\mathsf{NP} \cup \mathsf{coNP} \subseteq \Theta_2^p \subseteq \Delta_2^p \subseteq \Sigma_2^p$.

2.2 Trajectoires préférées

Définition 1 *Un* scénario $\Sigma = \langle \varphi_1, ..., \varphi_N \rangle$ *est une séquence ordonnée de formules propositionnelles. Le* t^e *élément* (φ_t) *de* Σ *est noté* $\Sigma(t)$. $|\Sigma|$ *est la longueur de* Σ. *Un scénario est* cohérent *ssi* $\forall t \in 1..|\Sigma|$, $\Sigma(t)$ *est cohérent.* $\mathcal{S}$ *(resp.* $\mathcal{S}_N$*) est l'ensemble de tous les scénarios (resp. de tous les scénarios de longueur* N*). Pour* $\Sigma, \Sigma' \in \mathcal{S}_N$, *on écrit que* $\Sigma \models \Sigma'$ *ssi* $\forall t$, $\Sigma(t) \models \Sigma'(t)$, $\Sigma \equiv \Sigma'$ *ssi* $\forall t$, $\Sigma(t) \models \Sigma'(t)$ *et* $\Sigma'(t) \models \Sigma(t)$, *et on définit* $\Sigma \wedge \Sigma' = \langle \Sigma(t) \wedge \Sigma'(t) | t \in 1 \ldots N \rangle$.

La concaténation $\Sigma.\Sigma'$ *de deux scénarios* Σ *et* Σ' *est définie par* $(\Sigma.\Sigma')(t) = \left\{ \begin{array}{l} \Sigma(t) \text{ si } t \leq |\Sigma| \text{ et} \\ \Sigma'(t - |\Sigma|) \text{ si } t > |\Sigma| \end{array} \right\}$.

Un scénario Σ *tel que* $\Sigma(t) = \varphi$ *et* $\Sigma(t') = \top\ \forall t' \neq t$ *est*

3. U8 $(K_1 \vee K_2) \diamond \varphi \equiv (K_1 \diamond \varphi) \vee (K_2 \diamond \varphi)$ [11]

4. Un fluent est une proposition dont la valeur de vérité n'est pas figée, c'est-à-dire peut évoluer dans le temps.

abrégé en $[t]\varphi$.

Définition 2 *Une* trajectoire *de longueur N est une suite $\tau = \langle \tau(1), .., \tau(N) \rangle$ de modèles de $\mathcal{M}$. $TRAJ_N$ représente l'ensemble de toutes les trajectoires de longueur N. On dit que $\tau \in TRAJ_N$ satisfait $\Sigma \in \mathcal{S}_N$ (noté $\tau \models \Sigma$) ssi $\forall t \in 1..N$ on a $\tau(t) \models \Sigma(t)$, et on note $Traj(\Sigma) = \{\tau \in TRAJ_{|\Sigma|} | \tau \models \Sigma\}$. τ est* statique *ssi $\tau(1) = ... = \tau(N)$.*

L'extrapolation des croyances part d'un ensemble d'observations (provenant de capteurs, ou de dialogues avec d'autres agents, etc.), supposées fiables, constituant un scénario cohérent $\Sigma \in \mathcal{S}^5$. L'*extrapolation des croyances* consiste compléter Σ en utilisant des hypothèses de persistance par défaut. L'idée de l'extrapolation des croyances est que si rien ne dit le contraire, les fluents ne changent pas de valeur. On désire donc trouver les meilleures trajectoires satisfaisant les observations, c'est-à-dire les trajectoires les plus inertes possible. C'est pourquoi on introduit des relations de préférences sur les trajectoires :

Définition 3 *Une* relation de préférence *$\preceq_N$ est une relation réflexive et transitive sur $TRAJ_N$ (pas nécessairement totale). $\preceq$ représente une famille des relations de préférence $\preceq_N$ pour $N \in \mathbb{N}^{*\,6}$. On écrit $\tau \prec \tau'$ (se lit "τ est strictement préférée à τ'") pour ($\tau \preceq \tau'$ et non ($\tau' \preceq \tau$)), et $\tau \sim \tau'$ pour ($\tau \preceq \tau'$ et $\tau' \preceq \tau$). Si $X \subseteq TRAJ_N$, une trajectoire $\tau \in X$ est* minimale *pour $\preceq$ dans X ssi il n'y a pas de $\tau' \in X$ telle que $\tau' \prec \tau$.*

On définit un opérateur d'extrapolation basé sur une relation de préférence sur les trajectoires de la façon suivante :

Définition 4 *L'opérateur $\updownarrow_\preceq$: $\mathcal{S} \rightarrow \mathcal{S}$ induit par $\preceq$ qui transforme un scenario en un scenario extrapolé est défini par*

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

Nous allons maintenant donner plusieurs exemples d'opérateurs d'extrapolation. Nous définissons d'abord l'*ensemble des changements* $Ch(\tau)$ d'une trajectoire τ par
$Ch(\tau) = \{\langle f, t \rangle \,|\, f \in \mathcal{F}, t \in 2..|\tau|, \tau(t-1) \models \neg f \text{ et } \tau(t) \models f\}$[7].
On note aussi $Ch(\tau, t) = \{f \mid \langle f, t \rangle \in Ch(\tau)\}$ (l'ensemble des fluents qui deviennent vrai entre $t-1$ et t).
Pour toute $\tau \in TRAJ_N$ et tout $p \leq N-1$ tels que $\tau(p) = \tau(p+1)$ on définit $Skip(\tau, p) \in TRAJ_{N-1}$ par
$$Skip(\tau, p)(t) = \begin{cases} \tau(t) & \text{si } t \leq p-1 \\ \tau(t+1) & \text{si } t \in \{p, .., N-1\} \end{cases}$$

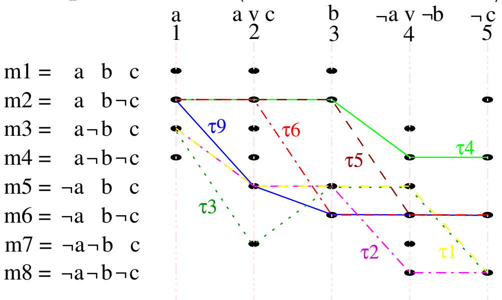
5. Pour simplifier, nous n'évoquons pas la possibilité d'exprimer des relations de dépendance entre fluents au moyen de lois statiques; cependant, il est possible (quoique non optimal en pratique) de donner à ces lois statiques le statut d'observations : à chaque instant, les croyances initiales seraient alors la conjonction des observations et des lois statiques.

6. Par la suite, nous omettrons l'indice et sans ambiguïté, nous écrirons $\tau \preceq \tau'$ au lieu de $\tau \preceq_N \tau'$, en supposant par défaut que τ et τ' ont la même longueur N (les trajectoires de longueur différentes ne sont jamais comparées).

7. Notons qu'une trajectoire peut être complètement définie par l'ensemble de ses changements et par $\tau(i)$ à un instant i donné (l'ensemble des changements ne suffit pas).

Cette trajectoire est la trajectoire initiale à laquelle on a enlevé un instant auquel aucun changement ne se produisait. Enfin, la *trajectoire regroupée en $p \leq N-1$* de τ, $\Phi_p(\tau) \in TRAJ_N$, est définie par :
- si $\forall f \in \mathcal{F}, f \in Ch(\tau, p) \Rightarrow \neg f \notin Ch(\tau, p+1)$ alors $\Phi_p(\tau)(p) = \tau(p+1)$ et $\forall t \neq p$, $\Phi_p(\tau)(t) = \tau(t)$;
- sinon, $\Phi_p(\tau) = \tau$.
La trajectoire obtenue est identique à la première à ceci près que l'on a regroupé les changements de p-1 à p et de p à p+1 en un seul changement de p-1 à p.

Exemple 1 *soit* $\Sigma = \langle a, a \vee c, b, \neg a \vee \neg b, \neg c \rangle$.



Considérons les onze trajectoires suivantes :

trajectoire τ	$Ch(\tau)$
$\tau_1 = \langle m3, m5, m5, m5, m8 \rangle$	$\{\langle \neg a, 2 \rangle, \langle b, 2 \rangle, \langle \neg b, 5 \rangle, \langle \neg c, 5 \rangle\}$
$\tau_2 = \langle m3, m5, m5, m8, m8 \rangle$	$\{\langle \neg a, 2 \rangle, \langle b, 2 \rangle, \langle \neg b, 4 \rangle, \langle \neg c, 4 \rangle\}$
$\tau_3 = \langle m3, m7, m5, m5, m8 \rangle$	$\{\langle \neg a, 2 \rangle, \langle b, 3 \rangle, \langle \neg b, 5 \rangle, \langle \neg c, 5 \rangle\}$
$\tau_4 = \langle m2, m2, m2, m4, m4 \rangle$	$\{\langle \neg b, 4 \rangle\}$
$\tau_5 = \langle m2, m2, m2, m6, m6 \rangle$	$\{\langle \neg a, 4 \rangle\}$
$\tau_6 = \langle m2, m2, m6, m6, m6 \rangle$	$\{\langle \neg a, 3 \rangle\}$
$\tau_7 = \langle m4, m4, m6, m6, m6 \rangle$	$\{\langle \neg a, 3 \rangle, \langle b, 3 \rangle\}$
$\tau_8 = \langle m1, m1, m6, m6, m6 \rangle$	$\{\langle \neg a, 3 \rangle, \langle \neg c, 3 \rangle\}$
$\tau_9 = \langle m1, m5, m6, m6, m6 \rangle$	$\{\langle \neg a, 2 \rangle, \langle \neg c, 3 \rangle\}$
$\tau_{10} = \langle m1, m5, m5, m6, m6 \rangle$	$\{\langle \neg a, 2 \rangle, \langle \neg c, 4 \rangle\}$
$\tau_{11} = \langle m1, m5, m5, m5, m6 \rangle$	$\{\langle \neg a, 2 \rangle, \langle \neg c, 5 \rangle\}$

τ_1 et τ_2 sont assez similaires, puisque les mêmes changements s'y produisent et dans le même ordre – τ_3 en est plus éloignée car, à l'instant 2, une seule variable change (et pas deux en même temps). Notons que $Skip(\tau_1, 2) = Skip(\tau_2, 4) = \langle m3, m5, m5, m8 \rangle$, et que le regroupement de la trajectoire τ_3 à l'instant 2 est $\Phi_2(\tau_3) = \tau_2$.

Donnons à présent quelques propriétés que peut posséder une relation de préférence entre trajectoires :
(I): $\preceq$ est dite *inertielle* ssi (1) pour deux trajectoires statiques quelconques de même longueur τ et τ' on a $\tau \sim \tau'$ et (2) pour toute trajectoire statique τ et non statique τ' (de même longueur) on a $\tau \prec \tau'$.
Cette propriété, très naturelle, est désormais requise pour toutes les relations de préférences. Les propriétés suivantes sont plus discutables :
(IRC): $\preceq$ est *insensible au regroupement des changements* ssi $\forall \tau_1, \tau_2 \in TRAJ_N$ et $\forall p \leq N-1$, $\tau_1 \preceq \tau_2$ implique $\Phi_p(\tau_1) \preceq \Phi_p(\tau_2)$. (IRC) exprime que deux ensembles de changements concernant des variables distinctes à des instants consécutifs peuvent être regroupés sans changer la relation de préférence.
(ISI): $\preceq$ est *insensible aux sauts d'instants* ssi $\forall \tau_1, \tau_2 \in TRAJ_N$ et p tels que $\tau_1(p) = \tau_1(p+1)$ et $\tau_2(p) = \tau_2(p+1)$, on a $\tau_1 \preceq \tau_2$ ssi $Skip(\tau_1, p) \preceq Skip(\tau_2, p)$.

(D): $\preceq$ satisfait la *décomposabilité* ssi $\forall \tau_1, \tau_2 \in TRAJ_p$, $\tau_3, \tau_4 \in TRAJ_{N-p}$ tels que $\tau_1(p) = \tau_3(1)$ et $\tau_2(p) = \tau_4(1)$, on a : $\tau_1 \preceq \tau_2$ et $\tau_3 \preceq \tau_4$ implique $\tau_1.\tau_3 \preceq \tau_2.\tau_4$.

Voici quelques exemples de relations de préférence :

- nombre de changements :
$\tau \preceq_{nc} \tau'$ ssi $|Ch(\tau)| \leq |Ch(\tau')|$

- nombre de changements par fluent :
$\tau \preceq_{ncf} \tau'$ ssi $\forall f \in \mathcal{F}, |Ch(\tau, f)| \leq |Ch(\tau', f)|$;

- nombre d'instants de changements :
$\tau \preceq_{nic} \tau'$ ssi
$|\{t \mid \exists f \text{ t.q. } \langle f, t\rangle\ Ch(\tau)\}| \leq |\{t \mid \exists f \text{ t.q. } \langle f, t\rangle\ Ch(\tau')\}|$;

- inclusion de l'ensemble des changements :
$\tau \preceq_{iec} \tau'$ ssi $Ch(\tau) \subseteq Ch(\tau')$;

- inclusion des fluents changeants : $\tau \preceq_{ifc} \tau'$ ssi
$\{f \in \mathcal{F} \mid \exists t, \langle f, t\rangle \in Ch(\tau)\} \subseteq \{f \in \mathcal{F} \mid \exists t, \langle f, t\rangle \in Ch(\tau')\}$

- pénalité : soit $k : \mathcal{F} \rightarrow \mathbb{N}^*$: $k(f)$ est la pénalité induite pour un changement de $\neg f$ en f, et soit $K(\tau) = \sum_t \sum_{\langle f,t\rangle \in Ch(\tau)} k(f)$.
$\tau \preceq^k_{cp} \tau'$ ssi $K(\tau) \leq K(\tau')$[8];

- chronologique : soit $\tau = \langle M_1..M_p\rangle$, $\tau' = \langle M'_1..M'_p\rangle$; on définit $\tau \preceq_{chr} \tau'$ ssi $p = 1$ ou $Ch(\tau, 1) \subset Ch(\tau', 1)$ ou ($Ch(\tau, 1) = Ch(\tau', 1)$ et $\langle M_2..M_p\rangle \preceq_{chr} \langle M'_2..M'_p\rangle$). Cette relation préfère les trajectoires dont les changements ont lieu le plus tard possible (comme dans [21]).

- anti-chronologique : avec les notations précédentes, $\tau \preceq_{achr} \tau'$ ssi $p = 1$ ou $Ch(\tau, p) \subset Ch(\tau', p)$ ou ($Ch(\tau, p) = Ch(\tau', p)$ et $\langle M_1..M_{p-1}\rangle \preceq_{achr} \langle M'_1..M'_{p-1}\rangle$). Cette relation préfère les trajectoires dont les changements ont lieu le plus tôt possible.

Notons que $\preceq_{iec} \subset \preceq_{ncf} \subset \preceq_{nc}$, $\preceq_{iec} \subset \preceq_{ifc}$, $\preceq_{iec} \subset \preceq_{nic}$, $\preceq_{iec} \subset \preceq_{chr}$ et $\preceq_{iec} \subset \preceq_{achr}$.

Exemple 2 *Les trajectoires satisfaisant Σ et minimales pour **nc**, **ncf** et **ifc** sont τ_4, τ_5 et τ_6. **nic** admet onze trajectoires minimales dont par exemple τ_7, τ_8, τ_4, τ_5 et τ_6. **iec** différencie les instants auxquels ont lieu les changements et admet six trajectoires minimales τ_4, τ_5, τ_6, τ_9, τ_{10} et τ_{11}. τ_7, où les fluents qui changent sont les mêmes que pour τ_9 mais à des instants différents, n'est pas préférée (car $\tau_6 \preceq_{iec} \tau_7$) : $\preceq_{iec}$ tient compte de l'identité des instants. Les trajectoires minimales pour **chr** sont τ_4 et τ_5 ($\tau_4 \preceq_{chr} \tau_6$ ainsi que $\tau_5 \preceq_{chr} \tau_6$ puisque pour la minimisation chronologique il est préférable que les changements se produisent le plus tard possible). τ_6 et τ_9 sont les trajectoires minimales pour **achr**.*
En abrégeant $\updownarrow_{\preceq_x}$ en $\updownarrow_x$, on a :
$\Sigma\updownarrow_{nc} = \Sigma\updownarrow_{ncf} = \Sigma\updownarrow_{ifc} =$
$\langle a \wedge b \wedge \neg c, a \wedge b \wedge \neg c, b \wedge \neg c, (a \oplus b) \wedge \neg c, (a \oplus b) \wedge \neg c\rangle$;
$\Sigma\updownarrow_{nic} = \langle a, a, b \wedge (a \vee \neg c), (\neg a \vee \neg b) \wedge \neg c, (\neg a \vee \neg b) \wedge \neg c\rangle$;
$\Sigma\updownarrow_{iec} = \langle a \wedge b, (a \oplus c) \wedge b, (\neg a \wedge b) \vee (a \wedge b \wedge \neg c), (\neg a \wedge b) \vee (a \wedge \neg b \wedge \neg c), (a \oplus b) \wedge \neg c\rangle$;
$\Sigma\updownarrow_{chr} = \langle a \wedge b \wedge c, a \wedge b \wedge c, a \wedge b \wedge c, (a \oplus b) \wedge \neg c, (a \oplus b) \wedge \neg c\rangle$;
$\Sigma\updownarrow_{achr} = \langle a \wedge b, (a \oplus c) \wedge b, \neg a \wedge b \wedge \neg c, \neg a \wedge b \wedge \neg c, \neg a \wedge b \wedge \neg c\rangle$.

Il peut arriver que des fluents ne changent pas de façon indépendante les uns des autres, en particulier quand les changements sont dus à des *événements*. Un *événement* (déterministe) ε est une fonction de $\mathcal{M}$ vers $\mathcal{M}$, où $\varepsilon(M)$ est le résultat de ε quand l'état initial est M[9]. L'*événement vide* $\varepsilon_\emptyset$ est l'identité. Chaque événement est associé à une pénalité $k(\varepsilon) \geq 0$ avec $k(\varepsilon) = 0$ ssi $\varepsilon = \varepsilon_\emptyset$, et pour chaque séquence d'événements $\theta = \varepsilon_1 \circ ... \circ \varepsilon_p$ on pose $k(\theta) = \sum_{i \in 1..p} k(\varepsilon_i)$. Étant donné un ensemble d'événements E et deux modèles M, M', soit $k^E(M, M') = \min_{\theta \mid \theta(M) = M'} k(\theta)$ (et $= +\infty$ si une telle séquence n'existe pas) et $K^E(\tau) = \sum_{t=0..N-1} k^E(\tau(t), \tau(t+1))$. On définit maintenant la relation de préférence basée sur les événements :

- pénalité d'événement :
$\tau \preceq^E_e \tau'$ ssi $K^E(\tau) \leq K^E(\tau')$

2.3 Propriétés de l'extrapolation

Définition 5 *Considérons les propriétés suivantes qu'un opérateur $\updownarrow$: $\mathcal{S} \rightarrow \mathcal{S}$ qui transforme un scenario en un scenario extrapolé peut satisfaire pour tout scénario cohérent Σ :*

E1 $\Sigma\updownarrow \models \Sigma$;
E2 $\Sigma\updownarrow$ *est cohérent;*
E3 *si* $\Sigma \equiv \Sigma'$ *alors* $\Sigma\updownarrow \equiv \Sigma'\updownarrow$;
E4 $\Sigma\updownarrow(t) \wedge \varphi \models (\Sigma \wedge [t]\varphi)\updownarrow(t)$;
E5 $\bigwedge_{t' \in 0..N} \Sigma(t')$ *cohérent* $\Rightarrow \forall t,\ \Sigma\updownarrow(t) = \bigwedge_{t' \in 0..N} \Sigma(t')$;
E6 $\Sigma\updownarrow(t) \wedge \varphi$ *cohérent* $\Rightarrow (\Sigma \wedge [t]\varphi)\updownarrow(t) \models \Sigma\updownarrow(t) \wedge \varphi$;
E7 *si* $\varphi \models \psi$ *ou* $\psi \models \varphi$ *alors* $\forall t \neq |\Sigma| + 1, |\Sigma| + 2$ *on a* $(\Sigma.\varphi.\psi.\Sigma')\updownarrow(t) \equiv (\Sigma.\varphi \wedge \psi.\varphi \wedge \psi.\Sigma')\updownarrow(t)$;
E8 $\forall t \neq |\Sigma|+1, (\Sigma.\top.\Sigma')\updownarrow(t) \equiv \begin{cases} (\Sigma.\Sigma')\updownarrow(t) & si\ t \leq |\Sigma| + 1 \\ (\Sigma.\Sigma')\updownarrow(t-1) & si\ t > |\Sigma| + 1 \end{cases}$
E9 *si* $(\Sigma.\varphi)\updownarrow(|\Sigma| + 1) \equiv (\varphi.\Sigma')\updownarrow(1)$ *alors*
$(\Sigma.\varphi.\Sigma')\updownarrow(t) \equiv \begin{cases} (\Sigma.\varphi)\updownarrow(t) & si\ t \leq |\Sigma| + 1 \\ (\varphi.\Sigma')\updownarrow(t - |\Sigma|) & si\ t > |\Sigma| + 1 \end{cases}$

E1 signifie que les observations sont fiables. E2 et E3 sont naturels. E4 signifie que l'extrapolation est monotone quand on ajoute une nouvelle observation cohérente avec l'extrapolation des observations initiales, et la conjonction de E4 et E6 signifie que la prise en compte d'une nouvelle observation cohérente avec $\Sigma\updownarrow$ avant ou après extrapolation donne des résultats équivalents. E5 signifie que si on peut supposer de façon cohérente que rien ne change, alors cette supposition peut être faite et ainsi la connaissance extrapolée à chaque instant est la conjonction de toutes les observations. Notons que E1, E2 et E4 impliquent $\Sigma\updownarrow\updownarrow \equiv \Sigma\updownarrow$, (idempotence de l'extrapolation). Un opérateur d'extrapolation satisfaisant E1-E5[10] (qui semblent indiscutables) sera dit *basique*.

E7 exprime que si une observation est suivie ou précédée d'une observation plus forte, alors les croyances extrapolées aux autres instants seraient les mêmes que si on avait eu plusieurs fois la plus forte[11]. E8 signifie qu'une ob-

8. Notons que l'on n'a pas nécessairement $k(f) = k(\neg f)$. Par ailleurs, si $k(f) = 1$ pour tout f, on retrouve la relation $\preceq_{nc}$.

9. Les événements n'ont pas besoin d'être représentés explicitement. On peut utiliser par exemple des listes STRIPS comme pour les actions.

10. c'est-à-dire E1, E2, E3, E4 et E5, bien entendu.

11. Attention, ce postulat considère la persistance vers le futur mais aussi vers le passé : si une observation est vraie en un instant, alors, par défaut, elle est vraie après et aussi avant.

servation vide n'a pas d'influence sur l'extrapolation des croyances; E9 est une propriété markovienne qui signifie qu'un problème d'extrapolation peut (sous certaines conditions) être décomposé en sous-problèmes indépendants.

Proposition 1
(1) Tout opérateur $\updownarrow_{\preceq}$ *induit par une relation de préférence inertielle* $\preceq$ *satisfait E1-E5.*
(2) Tout opérateur $\updownarrow_{\preceq}$ *induit par une relation de préférence inertielle totale* $\preceq$ *satisfait E1-E6.* [12]

On peut montrer que tout opérateur induit à partir d'une relation de préférence inertielle satisfaisant en outre IRC (respectivement ISI, D) satisfait E7 (respectivement E8,E9).

Proposition 2 *Les opérateurs* $\updownarrow_{nc}$, $\updownarrow_{cp}^{k}$ *et* $\updownarrow_{nic}$ *satisfont E1-E9. Les opérateurs* $\updownarrow_{iec}$ *et* $\updownarrow_{ncf}$ *satisfont E1-E5 et E7-E9. Les opérateurs* $\updownarrow_{chr}$ *et* $\updownarrow_{achr}$ *satisfont E1-E5, E8 et E9. L'opérateur* $\updownarrow_{ifc}$ *satisfait E1-E5, E7 et E8. L'opérateur* $\updownarrow_{e}^{E}$ *satisfait E1-E6, E8 et E9.* [13]

Si $\Sigma = \langle \Sigma(1), .., \Sigma(N) \rangle$, soit $Renverse(\Sigma) = \langle \Sigma(N), .., \Sigma(1) \rangle$, et définissons la propriété **R** ainsi :

R $Renverse(\Sigma \updownarrow) = (Renverse(\Sigma)) \updownarrow$.

On peut montrer que $\updownarrow_{\preceq}$ satisfait R ssi $\preceq$ est tel que pour tout τ, τ' de même longueur N on a $\tau \preceq \tau'$ ssi $Renverse(\tau) \preceq Renverse(\tau')$ (où $Renverse(\tau)(t) = \tau(N-t)$).

Proposition 3 *Les opérateurs d'extrapolation* $\updownarrow_{iec}$, $\updownarrow_{ncf}$, $\updownarrow_{nc}$, $\updownarrow_{nic}$ *et* $\updownarrow_{ifc}$ *satisfont (R).* $\updownarrow_{cp}^{k}$ *satisfait (R) ssi* $\forall f \in \mathcal{F}$ *on a* $k(f) = k(\neg f)$. $\updownarrow_{chr}$ *et* $\updownarrow_{achr}$ *ne satisfont pas (R).* $\updownarrow_{ep}^{E}$ *satisfait (R) ssi* $\forall M, M'$ *on a* $k^{E}(M, M') = k^{E}(M', M)$.

3 Complexité et calcul de la complétion extrapolatoire

Identifions tout d'abord la complexité des opérateurs d'extrapolation précédents.

Proposition 4
EXTRA$_x$ *est le problème de décision suivant : étant donnés* Σ, ψ *et* t, *décider si* $\Sigma \updownarrow_x (t) \models \psi$ *est vrai.*

- EXTRA$_{iec}$, EXTRA$_{ifc}$, EXTRA$_{ncf}$, EXTRA$_{chr}$ *et* EXTRA$_{achr}$ *sont* Π_2^p*-complets;*
- EXTRA$_{cp}^{k}$ *est dans* Δ_2^{P}*;*
- EXTRA$_{nc}$ *et* EXTRA$_{nic}$ *sont dans* $\Delta_2^{\mathsf{P}}[\mathcal{O}(\log n)]$.

Ces résultats ne sont pas surprenants : en effet, la plupart des problèmes de changement des croyances sont à ces niveaux de complexité (voir par exemple [9]).
Pour le reste de ce paragraphe, on utilise $\preceq_{iec}$ *comme relation de préférence.*

Définition 6 *Un ensemble de changements* γ [14] *est une* explication minimale *pour* Σ *ssi il existe une trajectoire* $\preceq_{iec}$*-préférée* τ *de* Σ *telle que* $Ch(\tau) = \gamma$.

L'ensemble de toutes les explications minimales est très riche : même s'il ne détermine pas toujours les trajectoires préférées dans le cas général (parce que plusieurs trajectoires peuvent avoir le même ensemble de changements), il le permet dès qu'il existe un t_i tel que $\Sigma(t_i)$ est *complet* (c'est-à-dire, n'admet qu'un seul modèle M_i), et dans ce cas, à partir de l'ensemble des explications minimales et de M_t on peut décider en temps polynomial si $\Sigma \updownarrow (t) \models \psi$ est vrai pour tout t et ψ.
Étant donné un scénario Σ (de longueur N), le nombre d'explications minimales peut être très grand, même quand le nombre de changements est petit :

Exemple 3 $\Sigma = \langle a \wedge b \wedge c \wedge d, \top, \neg b, d, \neg a, \top, b, \neg c \vee \neg d \rangle$ *a exactement* $4 \times 2 \times 4 \times (7+4) = 352$ *explications minimales (et autant de trajectoires minimales). Ces 352 explications minimales peuvent être exprimées de façon plus compacte par :* γ *est une explication minimale pour* Σ *ssi elle contient*
- $\langle \neg a, t \rangle$ *pour (exactement) un seul instant* $t \in \{2..5\}$ *(signifiant que* a *devient faux exactement une fois, à l'instant 2, 3, 4 ou 5 – car a est vrai à l'instant 1 et faux à l'instant 5);*
- $\langle \neg b, t \rangle$ *pour un* $t \in \{2, 3\}$, *et* $\langle b, t \rangle$ *pour un* $t \in \{4..7\}$*;*
- soit $\langle \neg c, t \rangle$ *pour un* $t \in \{2..8\}$, *soit* $\langle \neg d, t \rangle$ *pour un* $t \in \{5..8\}$*; et rien d'autre.*

Pour définir cette notion simple, nous introduisons la notion d'*ensemble compact de changements*.

Définition 7 *Un* ensemble compact de changements Γ *est un ensemble* $\{\langle f_i, t_i^-, t_i^+ \rangle \mid i = 1..p\}$ *où les* f_i *sont des fluents et les* (t_i^-, t_i^+) *des paires d'instants tels que* $0 < t_i^- \leq t_i^+ \leq N$. *L'ensemble des changements couverts par* Γ *est* $Cov(\Gamma) = \{\langle f_1, t_1 \rangle, ..., \langle f_p, t_p \rangle \mid \forall i = 1..p, t_i^- \leq t_i \leq t_i^+\}$.

Intuitivement, un ensemble compact de changements est constitué de la disjonction de tous les ensembles de changements qu'il couvre.

Définition 8 Γ explique Σ *ssi tout* $\gamma \in Cov(\Gamma)$ *explique* Σ, *on dit que* Γ explique minimalement Σ *ssi* $\forall \gamma \in Cov(\Gamma)$, γ *est une explication minimale pour* Σ *et que* Γ *est une* explication minimale maximalement compacte (MCME) *pour* Σ *ssi (i)* Γ *explique minimalement* Σ *et (ii) il n'existe pas* Γ' *expliquant minimalement* Σ *telle que* $Cov(\Gamma') \subset Cov(\Gamma)$.

12. La réciproque de (2) est probablement vraie, *i.e.*, un opérateur $\updownarrow$ satisfait E1-E6 ssi il existe une relation de préférence totale inertielle $\preceq$ telle que $\updownarrow = \updownarrow_{\preceq}$. La réciproque de (1) n'est pas vraie : il manque un postulat, qui serait une forme plus faible de E6, que nous ne mentionnons pas du fait du manque de place et parce qu'il est compliqué et essentiellement technique. Voir aussi Section 4.

13. On a donc des exemples d'opérateurs d'extrapolation violant E6 ou E7 ou E9. Un exemple d'opérateur d'extrapolation ne satisfaisant pas E8 est l'opérateur – certes étrange – basé sur la relation $\tau \preceq \tau'$ ssi $\forall t = 2...N-1$, $\{f | \langle f, t \rangle \in Ch(\tau)$ ou $\langle f, t+1 \rangle \in Ch(\tau)\} \subseteq \{f | \langle f, t \rangle \in Ch(\tau')$ ou $\langle f, t+1 \rangle \in Ch(\tau')\}$.

14. c'est-à-dire, une liste de paires $\langle f, t \rangle$ – voir le paragraphe 2.2.

L'ensemble des MCMEs pour Σ regroupe les explications minimales d'une façon compacte et n'est pas redondant : en effet, on peut montrer facilement que pour toute explication minimale γ pour Σ il existe un et un seul MCME Γ pour Σ couvrant γ.
Dans l'exemple précédent, il y a deux explications minimales et donc deux MCMEs :
$\{\langle\neg a,2,5\rangle,\langle\neg b,2,3\rangle,\langle b,4,7\rangle,\langle\neg c,2,8\rangle\}$
et $\{\langle\neg a,2,5\rangle,\langle\neg b,2,3\rangle,\langle b,4,7\rangle,\langle\neg d,5,8\rangle\}$.

Proposition 5 *Si $\forall t$, $\Sigma(t)$ est une conjonction cohérente de fluents, alors il n'existe qu'une seule explication minimale maximalement compacte pour Σ.*

Les explications compactes minimales peuvent être calculées en ajoutant à Σ un ensemble maximal d'hypothèses de persistance exprimant que chaque fluent f persiste par défaut pendant les intervalles séparant deux observations consécutives "pertinentes" pour f. Cette notion de pertinence est représentée de façon simple (voir [14]): On dit qu'une observation φ est *pertinente* pour une variable propositionnelle v, noté $Rel(\varphi,f)$, ssi il n'y a pas de formule équivalente à φ dans laquelle v n'apparaît pas du tout[15]. Ainsi, pour une variable v, soit $R_\Sigma(v)=\{t\in 1..N \text{ s.t. } Rel(\Sigma(t),v)$; et l'ensemble des *intervalles pertinents* $RI_\Sigma(v)$ pour v par rapport à Σ est défini par $RI_\Sigma(v)=\{[t,t']\,|t\in R_\Sigma(v),t'\in R_\Sigma(v),[t+1,t'-1]\cap R_\Sigma(v)=\emptyset\}$. La signification intuitive des intervalles pertinents est qu'il est raisonnable de ne considérer les hypothèses de persistance portant sur une variable donnée que sur ces intervalles seulement.[16] On indexe toutes les variables par l'instant auquel elle se réfère (comme dans [10]) : si $v\in\mathcal{V}$ et $t\in 1..N$ alors v_t est vraie ssi v est vraie à l'instant t. Si ϕ est une formule construite sur $\mathcal{V}$ alors ϕ_t est obtenue à partir de ϕ en remplaçant chaque variable v de ϕ par v_t. Soit $\Sigma^*=\{\Sigma(t)_t\mid t=1..N\}$ et alors l'ensemble des *hypothèses de persistance* pour Σ est $P_\Sigma=\bigcup_{\substack{v\in\mathcal{V}\\ [t,t']\in RI_\Sigma(v)}}\{v_t\rightarrow v_{t'},\neg v_t\rightarrow\neg v_{t'}\}$. Définissons les *P-explications* pour $\langle P_\Sigma,\Sigma^*\rangle$ comme les sous-ensembles minimaux π de P_Σ tels que $P_\Sigma\setminus\pi$ est cohérent avec Σ^*. Chaque P-explication π peut être identifiée par un ensemble de changements compacts $\Gamma(\pi)$ de façon triviale ; par exemple, $\pi=\{a_4\rightarrow a_7,\neg c_5\rightarrow\neg c_6\}$ donne $\Gamma(\pi)=\{\langle\neg a,5,7\rangle,\langle c,6,6\rangle\}$, c'est-à-dire que a devient faux une fois entre 4 et 7 (c'est-à-dire, à 5, 6, ou 7) et c devient vrai en 6. On a donc :

Proposition 6 *(1) Pour toute P-explication π pour $\langle P_\Sigma,\Sigma^*\rangle$, chaque ensemble de changements $\gamma\in\Gamma(\pi)$ est une explication minimale pour Σ[17] ; (2) chaque explication minimale pour Σ est couverte par exactement une P-explication, c'est-à-dire qu'il existe une et une seule π telle que $\gamma\in\Gamma(\pi)$.*

Ainsi, un ensemble couvrant d'explications compactes peut être calculé en utilisant un algorithme abductif. Ce dernier résultat établit un lien avec la "base revision" [18]:

Proposition 7
$\Sigma\updownarrow(t)\models\psi$ (avec $\preceq_{iec}$) ssi $\psi_t\in P_\Sigma\star\Sigma^*$, où $\star$ est l'opérateur de "base revision" (sans priorités) de [18].

Ce résulat (similaire à la Proposition 1 de [5]) montre que l'extrapolation (avec $\preceq_{iec}$) peut être traduite en un problème de révision par l'intermédiaire d'une indexation des variables. En fait, *les croyances initiales que les fluents persistent sont révisées par les observations.* Ce résultat introduit la section suivante dans laquelle on montre que plus généralement, il existe des liens forts entre extrapolation et révision.

4 Extrapolation et révision

La révision des croyances traite généralement de mondes statiques, et par conséquent travaille avec des formules concernant le même monde au même instant. Toutefois, comme on peut le lire dans [10], "ce qui est important pour la révision n'est pas que le monde soit statique, mais que les propositions utilisées pour décrire le monde soient statiques". Ceci explique pourquoi, lorsqu'on indexe les variables propositionnelles par des instants comme à la Proposition 7, on réussit à capturer un processus dynamique dans un cadre de révision des croyances. Ce procédé n'est pas nouveau : ainsi, [5] proposent une traduction des opérateurs de mise-à-jour à partir de transitions en un opérateur de révision, moyennant un enrichissement du langage propositionnel identique à celui que nous avons donné. Ainsi, un processus d'extrapolation est équivalent à une *révision de la croyance initiale en la persistance des fluents par les observations*[18]

Un premier lien apparaît lorsqu'on considère des séquences de seulement *deux* observations. Des opérateurs d'extrapolation des séquences de longueur 2 sont appelées des opérateurs de *2-extrapolation*.

À partir de $\updownarrow$: $\mathcal{L}^2\rightarrow\mathcal{L}^2$, définissons l'opérateur de changement $\star=f(\updownarrow):\mathcal{L}^2\rightarrow\mathcal{L}^2$ par $KB\star\alpha=\langle B,\alpha\rangle\updownarrow(2)$. Inversement, à partir d'un opérateur de changement de

15. ou de façon équivalente, φ est *non*-pertinente pour v ssi $\varphi_{v\leftarrow\top}$, $\varphi_{v\leftarrow\bot}$ et φ sont logiquement équivalentes (où $\varphi_{v\leftarrow\top}$ (resp. $\varphi_{v\leftarrow\bot}$) est la formule φ dans laquelle on a remplacée la variable v par $\top$ (resp. $\bot$)).

16. Cette méthode pourra être améliorée par la suite en distinguant les observations *positives* des *négatives*, mais nous n'en parlons pas par manque de place.

17. Notons que les explications compactes Γ calculées selon la méthode précédente ne sont pas toujours *maximalement* compactes, comme on peut le voir dans l'exemple suivant : $\Sigma=\langle a,c,a\vee c,b,\neg a\vee b,\neg a\rangle$ donne les explications compactes minimales $\{\langle\neg a,4,5\rangle\}$ et $\langle\neg a,6,6\rangle\}$ qui ne sont pas *maximalement* compactes puisque leur union $\{\langle\neg,4,6\rangle\}$ est une MCME. Si l'on souhaite obtenir les MCMEs pour Σ on doit définir une procédure de compaction dont la description précise sort du cadre de cet article.

18. Cette traduction peut s'étendre à des cadres plus complexes où le monde n'est plus simplement inerte mais a une dynamique propre. Voir [5].

croyances $\star$, définissons l'opérateur $\updownarrow = g(\star)$ par [19] [20]

$$\langle \Sigma(1), \Sigma(2) \rangle \updownarrow (2) = \Sigma(1) \star \Sigma(2)$$

et

$$\langle \Sigma(1), \Sigma(2) \rangle \updownarrow (1) = \Sigma(2) \star \Sigma(1)$$

Proposition 8

1. *$\updownarrow$: $\mathcal{L}^2 \to \mathcal{L}^2$ satisfait E1-E6 si et seulement si $f(\updownarrow)$ est un opérateur de révision au sens d'AGM;*

2. *$\star$ est un opérateur de révision au sens d'AGM si et seulement si $g(\star)$ est un opérateur de 2-extrapolation satisfaisant E1-E6.*

Ce résultat nous éclaire sur le sens véritable de l'extrapolation de croyances. En supposant que $\preceq$ est total et que $N = 2$, les croyances extrapolées à l'instant $t = 2$ consistent en la révision des croyances à $t = 1$ (dont on croit a priori qu'elles ont persisté jusqu'à $t = 2$) par les observations à $t = 2$ (tenues pour certaines). De manière symétrique, les croyances extrapolées à l'instant $t = 1$ consistent en la révision des croyances à $t = 2$ (dont on croit a priori qu'elles étaient déjà vraies à $t = 1$) – "persistance vers le passé"– par les observations à $t = 1$.

Proposition 9 *Soit $\star$ un opérateur de révision au sens de AGM défini sur un langage propositionnel généré par un ensemble fini de variables $\mathcal{V} = \{v_1, .., v_n\}$. Indexons les variables par des instants comme en Section 3. Soit $PERS = \bigwedge_{\substack{v \in \mathcal{V} \\ t=1..N-1}} v_t \leftrightarrow v_{t+1}$ et $\Sigma^* = \{\Sigma(t)_t \mid t = 1..N\}$. Alors, l'opérateur $E : \Sigma \mapsto E(\Sigma)(t) = \psi_t$, où ψ_t est la formule logique la plus forte (au sens de la conséquence logique) ne mentionnant que des variables indexées par t et telle que $PERS \star \Sigma^* \models \psi_t$, est un opérateur d'extrapolation satisfaisant E1-E6.*

Ainsi, on peut induire un opérateur d'extrapolation à partir de n'importe quel opérateur de révision . En fait on pourrait montrer que les postulats E1-E6 correspondent à une reformulation des postulats d'AGM de manière à prendre en compte les différents instants; toutefois, ces postulats E1-E6 ne concernent pas véritablement l'aspect temporel du processus, d'où le besoin d'ajouter les postulats plus spécifiquement liés au temps E7-E9 qui nous ramènent à des processus Markoviens et qui stipulent l'indépendance des changements. Par conséquent, dès que des postulats autres que E1-E6 sont requis, l'extrapolation est "plus spécifique" que la révision au sens d'AGM.
Une autre question intéressante concerne les liens possibles entre l'extrapolation et la *révision itérée*. Tout particulièrement, la révision itérée au sens de Lehmann [16] opère sur des séquences de formules cohérentes [21], comme l'extrapolation. Comme le résultat de la révision itérée au sens de Lehmann est l'ensemble des croyances concernant l'instant final seulement, ce que nous pouvons comparer, étant donné un scénario Σ, est l'ensemble de croyances $[\Sigma]$ résultant de Σ dans [16] avec la projection de l'extrapolation de Σ sur le dernier instant, c'est-à-dire $\Sigma \updownarrow (N)$ (où $N = |\Sigma|$).

Proposition 10 *Soit $\updownarrow$ un opérateur d'extrapolation basique. L'opérateur $\Sigma \mapsto \Sigma \updownarrow (N)$ satisfait les postulats I1, I2, I3, I5 et I6 de [16]* [22].

Par contre, I4 n'est généralement pas satisfait [23] et I7 n'est jamais satisfaite. Ainsi, Lehmann a raison lorsqu'il dit que pour un opérateur de changement des croyances s'appliquant à des situations évolutives, "on doit probablement accepter tous les postulats [...] sauf I4 et I7" [16].
La différence entre extrapolation et révision itérée apparait clairement sur l'exemple suivant : soit $\Sigma = \langle a \to b, a, \neg a \rangle$; pour tout opérateur d'extrapolation raisonnable on a $\Sigma \updownarrow = \langle a \wedge b, a \wedge b, \neg a \wedge b \rangle$ (le changement de a en $\neg a$ entre 2 et 3 étant certain, la trajectoire préférée est celle où il n'y a pas d'autres changements). Or, l'opérateur de révision itérée au sens de Lehmann, de même que la plupart des opérateurs de révision itérée sur des états épistémiques (en prenant comme état épistémique initial l'état d'ignorance totale), en particulier ceux de [13] et [1], conduisent à $[a \to b, a, \neg a] = \neg a$. La raison qui est au coeur de cette différence est que la révision itérée concerne des informations, certes obtenues à différents instants, mais concernant *un monde statique* : ceci explique pourquoi, une fois que la nouvelle information $\neg a$ a "annihilé" la précédente en raison du principe de la primauté forte de la nouvelle information [2] [13], on a supprimé de ce fait les raisons de croire en b [24].

5 Extrapolation et mise-à-jour

Lorsque $N = 2$, la projection de l'extrapolation de Σ à l'instant 2 peut être réexprimée ainsi : à partir d'un ensemble de croyances $K = \Sigma(1)$ à l'instant 1 et d'une observation $\alpha = \Sigma(2)$ à l'instant 2, on calcule un ensemble

19. Puisque le langage est généré par un nombre fini de variables propositionnelles, on considère comme dans [12] que les opérateurs de révision agissent sur des formules propositionnelles plutôt que sur des ensembles de croyances.

20. Il est clair que $f(g(\star)) = \star$ (toutefois, $f(g(\updownarrow)) = \updownarrow$ n'est pas nécessairement vrai).

21. Les autres approches de la révision itérée agissent sur des états épistémiques plutôt que sur des séquences de formules, et la comparaison avec l'extrapolation est alors plus complexe à établir.

22. σ étant une séquence de formules, $[\sigma]$ l'ensemble de croyances qu'il induit, et "." dénotant la concaténation :
I1 $[\sigma]$ est cohérent; **I2** $a \in [\sigma.a]$; **I3** si $b \in [\sigma.a]$ alors $a \to b \in [\sigma]$; **I4** si $a \in [\sigma]$ alors $[\sigma.\sigma'] = [\sigma.a.\sigma']$; **I5** si $b \models a$ alors $[\sigma.a.b.\sigma'] = [\sigma.a.b.\sigma']$; **I6** si $\neg b \notin [\sigma.a]$ alors $[\sigma.a.b.\sigma'] = [\sigma.a.a \wedge b.\sigma']$; **I7** $[\sigma.b.\neg b] \subseteq Cn([\sigma], b)$.

23. Il l'est si et seulement si $\preceq$ satisfait la propriété suivante : pour tout $\Sigma = \langle \varphi_1, ...\varphi_N \rangle$ et tout $\tau = \langle \tau(1)..\tau(N) \rangle, \tau \in Min(Traj(\Sigma), \preceq)$ implique $\langle \tau(1)..\tau(N-1) \rangle \in Min(Traj(\langle \varphi_1, ..., \varphi_{N-1} \rangle), \preceq)$. Seuls les opérateurs chronologiques tels que $\updownarrow_{chr}$ satisfont cette propriété.

24. Ce principe de primauté forte induit une forte "directivité" de la révision itérée, qui est incompatible avec la propriété R, satisfaite par la plupart des opérateurs d'extrapolation.

de croyances complété à l'instant 2. Cela peut paraître similaire à la mise-à-jour de croyances, puisque dans la littérature de la mise-à-jour, on lit habituellement que la mise-à-jour "intègre deux informations concernant deux instants différents". Toutefois, cette interprétation de la mise-à-jour est quelque peu ambiguë. En effet, la 2-extrapolation concerne des observations à différents instants avec des hypothèses de changement minimal et pourtant, *ce n'est pas* de la mise-à-jour. Ce dernier fait provient de façon claire du point 1 de la Proposition 8 (ou de E5): on sait que les postulats d'AGM sont incohérent avec le postulat U8 [11] dès que le langage contient au moins deux symboles propositionnels[25]:

Proposition 11 *Supposons que le langage contient au moins deux symboles propositionnels. Soit $N = 2$, $\updownarrow$ un opérateur de 2-extrapolation, et $\star$ l'opérateur de changement de croyances induit à partir de $\updownarrow$ comme en Section 4. Alors $\star$ ne peut pas satisfaire U8. Par conséquent, aucun opérateur de changement de croyances induit par un opérateur de 2-extrapolation n'est un opérateur de mise-à-jour.*

Ce résultat simple a des conséquences importantes. Il signifie que la mise-à-jour n'est pas pertinente pour raisonner avec des observations datées sur un monde évolutif (voir l'exemple introductif). Le point clé est le postulat U8 qui, en requérant que tous les modèles de l'ensemble initial de croyances soit mis-à-jour séparément, *nous empêche d'inférer des nouvelles croyances sur le passé à partir d'observations plus récentes* (ce qu'on nomme couramment *postdiction*). Cette inadéquation de la mise-à-jour n'est pas nouvelle (voir [4], [7] [6]); cependant, elle est rarement évoquée dans les articles sur la mise-à-jour, et même dans des articles récents on lit fréquemment l'idée – incorrecte – que la mise-à-jour de KB par α aurait comme but d'incorporer l'observation α (provenant par exemple de capteurs) reflètant un changement du monde. Cette dernière phrase serait correcte si au lieu de parler d'observations on parlait des *effets attendus d'une action* (le fait que ces effets puissent, dans certains cas, être observables – on devrait plutôt dire "vérifiables" – peut contribuer à entretenir la confusion).
Par conséquent, l'impossibilité, pour la mise-à-jour, d'inférer de nouvelles croyances à l'instant 1 à partir de croyances à l'instant 2 signifie que l'on est nécessairement dans l'une des situations suivantes : (1) on n'*observe rien* (à partir de capteurs, ou en communiquant avec d'autres agents etc.) à l'instant 2, ou du moins, rien qui ne soit *complètement prédictible* à partir de ce que l'on savait à l'instant 1 et des croyances a priori sur l'évolution du monde; (2) on observe des propriétés du monde à l'instant 2, mais ces observations sont soit *complètement décorrélées* du monde à l'instant 1 (ce qui implique l'abandon des hypothèses de persistance et de changement minimal). Des opérateurs de changement de croyances satisfaisant (2) constituent un cas dégénéré qui n'est pas très intéressant à étudier. Supposons donc (1) : par conséquent, si ce n'est pas une observation, la nouvelle formule α ne peut être qu'une *prédiction*, plus précisément, la projection des effets attendus d'un changement "explicite". Par *changement explicite*, nous voulons dire que quelque chose se produit, *et l'agent sait que cela se produit*, dont les résultats vont peut-être changer l'état du monde[26].
On voit donc que les questions essentielles sont *l'observabilité* (qu'observe-t-on du monde réel et à quels instants ?) et la *prédictabilité des changements*. L'extrapolation est adaptée aux observations et aux changements imprévus, tandis que la mise-à-jour est adapée aux changements attendus sans observations. Dans la classification de Sandewall's [20] des approches logiques du raisonnement sur un monde dynamique, l'extrapolation est adéquate pour la sous-classe sans actions de **K-IS** (connaissances correctes, inertie et surprises) tandis que la mise-à-jour est adéquate pour la classe $\mathsf{K_p}$-**IA** (pas d'observations après l'instant initial, inertie, actions à résultats alternatifs).
L'extrapolation et la mise-à-jour se complètent mutuellement, et si l'on veut être capable de raisonner à la fois avec des changements explicites et implicites, il faut alors intégrer les deux ce qui est relativement aisé. Nous évoquons comment cette intégration peut être faite. Un *scénario mixte* contient des observations et des changements explicites (correspondant à des effets attendus des actions); c'est donc une séquence de formules, chacune étant étiquetée soit par `observe` ou `update`, le premier élément de la séquence étant nécessairement du type `observe`. Par exemple, $\Sigma_0 = \langle$ `observe`: $a \vee b$, `update`: $\neg b$, `observe`: $\neg a$ $\rangle$ est un scénario mixte exprimant que les connaissances initiales sur le monde sont $a \vee b$, puis on fait une action dont l'effet attendu est $\neg b$, et, après son exécution, on observe $\neg a$.
Soit $\diamond$ un opérateur quelconque de mise-à-jour et $\preceq$ une relation de préférence comme en Section 2. Soit Σ un scénario mixte de longueur N, et $\tau = \langle M_0, ..., M_N \rangle$ une trajectoire. On dit que τ satisfait Σ par rapport à $\diamond$ si et seulement si (1) $\forall t$ tel que $\Sigma(t) = \texttt{observe:}\alpha$ on a $M_t \models \alpha$; (2) $\forall t$ tel que $\Sigma(t) = \texttt{update:}\alpha$ on a $M_t \in M_{t-1} \diamond \alpha$. Soit $Traj_\diamond(\Sigma)$ l'ensemble des trajectoires satisfaisant Σ par rapport à $\diamond$.
τ est une *trajectoire préférée* par rapport à $\diamond$ et $\preceq$ si et seulement si il n'existe pas de trajectoire $\tau' \in Traj_\diamond(\Sigma)$ telle que $\tau' \prec \tau$. Enfin, l'opérateur $\Updownarrow$ induit par $\preceq$ et $\diamond$ est défini par

$$Mod(\Sigma \Updownarrow (t)) = \{\tau(t) \mid \tau \in Min(Traj_\diamond(\Sigma), \preceq)\}$$

Les trajectoires préférées sont donc sélectionnées en deux étapes : (1) on élimine les trajectoires incompatibles avec

25. Plus précisément, U8 ne peut pas être satisfait en même temps que le (fort) postulat R2 si $KB \wedge \alpha$ est cohérent alors $KB \star \alpha \leftrightarrow KB \wedge \alpha$ en présence des postulats (plus inoffensifs) R1, R3 et R4.

26. C'est très différent d'une observation, qui peut éventuellement indiquer qu'un changement s'est produit mais de façon non explicite (on *infère* à partir de l'observation, quand elle est en conflit avec les croyances sur l'état précédent du monde, qu'un changement a eu lieu) : on pourrait parler de *changement implicite*.

au moins une observation ou un changement explicite; (2) au sein des trajectoires restantes, on minimise les changements imprévus[27].

Si l'on a choisi un opérateur $\diamond$ satisfaisant les postulats de Katsuno et Mendelzon et une relation de préférence $\preceq$ quelconque, avec le Σ_0 précédent on obtient $\Sigma_0 \Updownarrow= \langle a \vee b, \neg b, \neg a \wedge \neg b \rangle$.

Cette intégration de l'extrapolation et de la mise-à-jour permet de traiter des observations et des actions (déterministes ou non). Trois problèmes, cependant, ne sont pas traités : (a) les observations non fiables; (b) les actions à effets exceptionnels; (c) les événements complexes dont les effets attendus mènent à une mise-à-jour mais dont l'occurrence n'est pas connue a priori mais (abductivement) inférée des observations. Le troisième problème est traité (pour des séquences de longueur 2) par la mise-à-jour généralisée [4]. Le développement d'un cadre général intégrant mise-à-jour généralisée et extrapolation est en cours d'étude.

6 Liens avec d'autres travaux

Les liens entretenus par l'extrapolation avec la révision et la mise-à-jour ayant été présentés en Sections 4 et 5, parlons maintenant des liens avec d'autres travaux.

6.1 Berger, Lehmann et Schlechta

Dans [3], Berger, Lehmann et Schlechta ont proposé une approche pour le changement des croyances à base de trajectoires préférées, dans un esprit tout-à-fait similaire au nôtre. L'opérateur (générique) qu'ils proposent sont en effet des opérateurs d'extrapolation (même s'ils lui donnent le nom de "révision itérée" – ce qui est ambigu, pour les raisons expliquées en Section précédente). Plus exactement, un opérateur de Berger-Lehmann-Schlechta (BLS) est un opérateur d'extrapolation induit par une relation de préférence $\preceq$ qui vérifie la propriété (BLS) qui suit. Soit d'abord $h(\tau)$ la suite de mondes définie ainsi : $h(\tau)(1) = \tau(1)$; $h(\tau)(2) = \tau(i_2)$ où i_2 est le plus petit entier tel que $\tau(i_2) \neq \tau(1)$, s'il existe; pour $k \geq 2$, $h(\tau)(k+1) = \tau(i_{k+1})$ où i_{k+1} est le plus petit entier tel que $\tau(i_{k+1}) \neq \tau(i_k)$, s'il existe; et ainsi de suite jusqu'à ce qu'on ne puisse plus trouver de tel entier. Par exemple, pour les trajectoires de l'exemple 1 on a $h(\tau_1) = h(\tau_2) = \langle m_3, m_5, m_8 \rangle$ et $h(\tau_4) = \langle m_2, m_4 \rangle$. Une telle séquence est appelée une *histoire*. Étant données deux histoires h et h', on dit que h est une *sous-histoire* de h' si et seulement s'il existe une suite d'indices $1 \leq j_1 < j_2 < ... < j_k \leq$ tels que $h = \langle m_0, ..., m_n \rangle$ et $h' = \langle m_{j_0}, ..., m_{j_k} \rangle$. Voici enfin la propriété (BLS) :

(BLS) pour toutes trajectoires τ, τ' telles que $h(\tau)$ est une sous-histoire de $h(\tau')$, on a $\tau \preceq \tau'$.

On peut montrer que, parmi les relations de préférence que nous avons proposées, celles qui satisfont (BLS) sont $\preceq_{nc}$, $\preceq_{ncf}$, $\preceq_{nic}$, $\preceq_{ifc}$ et $\preceq_{cp}^k$; par contre, $\preceq_{iec}$, $\preceq_{chr}$ et $\preceq_{achr}$ ne satisfont pas (BLS).

6.2 Liberatore et Schaerf

Le système BReLS de Liberatore et Schaerf [17] intègre la révision, la mise-à-jour et la fusion de croyances. Il permet de traiter des observations datées. Deux sémantiques sont successivement considérées. La première, dite *sémantique des trajectoires*, s'apparente à l'extrapolation : en effet, en supposant qu'il n'y a pas plus d'une observation à chaque instant, on retrouve l'opérateur d'extrapolation $\preceq_{cp}$. La seconde, dite *sémantique point par point*, correspond à un processus de mise-à-jour itérée (mais est incompatible avec l'extrapolation à cause du postulat (U8) sous-jacent à cette sémantique). Curieusement, les auteurs ne considèrent pas l'éventualité de mélanger mises-à-jour et observations en utilisant les deux sémantiques concommitamment, comme nous l'avons brièvement fait en Section 5. En tout état de cause, leurs travaux et les nôtres (ces derniers s'attachant à étudier les propriétés logiques et calculatoires de différents opérateurs d'extrapolation) sont complémentaires.

6.3 Friedman et Halpern

Friedman et Halpern[10] définissent un cadre très général pour le changement de croyances basé sur les trajectoires (ou "runs") dans lequel le processus central est le *conditionnement*. Comme la révision et la mise-à-jour, on peut montrer que l'extrapolation entre dans leur cadre. Il resterait à déterminer précisément l'opération de conditionnement correspondant à l'extrapolation, ce qui est une piste de recherche intéressante.

6.4 Largouët et Cordier

Largouët et Cordier [15] proposent un cadre (probabiliste) pour le changement des croyances intégrant observations et évolutions attendues du monde ("expectations"), permettant, comme nous, de compléter un scénario initial par des hypothèses de persistance tant vers le futur que vers le passé. Leur processus de complétion opère ainsi :

soit $\Sigma = \langle obs(1), ..., obs(N) \rangle$ une séquence d'observations;

`pour` $t := 2$ `à` N `faire`

prédiction de $t-1$ à t; combinaison du résultat avec $obs(t)$;

`pour` $t := N-1$ `à 1 faire`

postdiction de t à $t-1$; combinaison du résultat avec $obs(t)$;

Ce n'est a priori pas aisé de comparer notre approche à la leur, en raison des hypothèses ontologiques qui diffèrent; cela devient néanmoins possible si on simplifie leur approche (1) en la "déprobabilisant" et (2) en supposant qu'il n'y a pas d'autre évolution attendue du monde que les lois de persistance par défaut (prédiction et postdiction se ramènent alors à persistance en avant et en arrière), et que les observations sont fiables. Soit alors une fonction de combinaison f telle que $f(\phi_1, ..., \phi_k) \models \phi_k$[28]. On définit alors l'opérateur $\Updownarrow_f$ par :

27. Notons que l'on peut avoir une observation *et* une mise-à-jour au même instant (sauf pour l'instant 1), ce qui signifie qu'une action ou un événement vient d'avoir lieu, et que l'observation est le feedback de cette action ou cet événement. De façon claire, l'observation *suit* la mise-à-jour et doit donc être prise en compte après. Cette hypothèse naturelle est également faite dans [4].

28. par exemple un opérateur de révision ou de révision itérée

$$\begin{aligned}\Sigma \downarrow\uparrow_f = \langle\ & f(obs(1), \ldots, obs(N), obs(N-1), \ldots, obs(1)), \\ & f(obs(1), \ldots, obs(N), obs(N-1), \ldots, obs(2)), \\ & \ldots, f(obs(1), \ldots, obs(N)) \qquad\qquad \rangle\end{aligned}$$

Soit $\Sigma = \langle \neg a, b, a \rangle$ et supposons les changements de a et b indépendants a priori. Après les phases prédiction-combinaison on a $\langle \neg a, \neg a \wedge b, a \wedge b \rangle$, puis après les phases de postdiction-combinaison on obtient $\langle \neg a \wedge b, a \wedge b, a \wedge b \rangle$, alors que la plupart des opérateurs d'extrapolation (sauf ceux qui sont similaires à $\updownarrow_{achr}$), donnent $\langle \neg a \wedge b, b, a \wedge b \rangle$: la différence entre les deux processus porte sur la croyance en $t = 2$. Le processus $\downarrow\uparrow_f$, en effectuant les phases de postdiction *après* celles de prédiction, donne la priorité à la persistance en arrière. Il s'agirait maintenant d'étudier de façon plus poussée les liens entretenus entre le processus de [15] et l'extrapolation.

6.5 Autres travaux

Dans des travaux précédents [8], nous avons considéré un modèle probabiliste des changements imprévus (qui peut être vu comme la contrepartie probabiliste de l'extrapolation), où une probabilité a priori est associée à chaque changement élémentaire, avec l'hypothèse d'indépendance entre changements élémentaires.

Une dernière ligne de recherche consisterait à poursuivre l'étude de la signification profonde des processus de révision, mise-à-jour et extrapolation, par exemple en situant l'extrapolation dans l'étude comparée entre filtre de Kalman et opérateurs de changement des cro-yances élaborée dans [7].

Références

[1] S. Benferhat, D. Dubois, and O. Papini. A sequential reversible belief revision method based on polynomials. In *AAAI-99*, pages 733–738, 1999.

[2] S. Benferhat, S. Konieczny, O. Papini, and R. Pino-Pérez. Révision itérée basée sur la primauté forte de la nouvelle information. In *Journées Nationales sur les Modèles de Raisonnement (JNMR'99)*, http://www.irit.fr/GDR-i3-ModRais/articlesJNMR.html.

[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] M.O. Cordier and J. Lang. Linking transition-based update and base revision. In *Lecture Notes in Artificial Intelligence 946 (Proc. of ECSQARU-95)*, pages 133–141, Fribourg, Switzerland, 1995.

[6] M.O. Cordier and C. Largouët. Combining observations and expectations: Application to the refinement of an image sequence. In *Proc. Workshop UAI*, 2000.

[7] C. Cossart and C. Tessier. Filtering vs. revision and update : let us debate! In *Proceedings of ECSQARU'99*, pages 116–127, 1999.

[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] T. Eiter and G. Gottlob. On the complexity of propositional knowledge base revision, updates and counterfactuals. *Artificial Intelligence*, 57:227–270, 1992.

[10] N. Friedman and J.Y. Halpern. Modeling beliefs in dynamic systems. part 2: revision and update. *JAIR*, 10:117–167, 1999.

[11] 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.

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

[13] S. Konieczny. *Sur la logique du changement : révision et fusion de bases de connaissances*. PhD thesis, Uniersité des SDciences et Technologies de Lille, 1999.

[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 *Workshop UAI : 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] B. Nebel. Base revision operations and schemes: Semantics, representation, and complexity. In *Proc. of ECAI'94*, pages 341–345.

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

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

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