A Stochastic Belief Change Framework with an Observation Stream and Defaults as Expired Observations

Gavin Rens[1]

Abstract. A framework for an agent to change its probabilistic beliefs after a stream of noisy observations is received is proposed. Observations which are no longer relevant, become default assumptions until overridden by newer, more prevalent observations. A distinction is made between background and foreground beliefs. Agent actions and environment events are distinguishable and form part of the agent model. It is left up to the agent designer to provide an environment model; a submodel of the agent model. An example of an environment model is provided in the paper, and an example scenario is based on it. Given the particular form of the agent model, several 'patterns of cognition' can be identified. An argument is made for four particular patterns.

1 MOTIVATION

My intention with this research is to design a framework with which an agent can deal with uncertainty about its observations and actions, and refresh its beliefs in a relatively sophisticated way.

Partially observable Markov decision processes (POMDPs) [1, 25, 23] are adequate for many stochastic domains, and they have the supporting theory to *update* agents' belief states due to a changing world. But POMDPs are lacking in two aspects with respect to intelligent agents, namely, (i) the ability to maintain and reason with background knowledge (besides the models inherent in POMDP structures) and (ii) the theory to *revise* beliefs due to information acquisition. Traditionally, *belief update* consists of bringing an agent's knowledge base up to date when the world described by the knowledge base changes, that is, it is a 'change-recording' operation, whereas *belief revision* is used when an agent obtains new information about a static world, that is, it is a 'knowledge-changing' operation [19]. I shall use the generic term *belief change* to including belief update and belief revision.

A different perspective on the proposed framework is from the area of classical belief change [12, 13, 17]. The belief change community has not given much attention to dealing with uncertainty, especially not stochastic uncertainty. Hence, integrating POMDP theory into belief change methods could be beneficial. And besides bringing probability theory, POMDPs also bring decision theory, that is, the theory for reasoning about actions and their utility.

However, I go one or two steps farther than a straightforward integration of POMDPs and belief change theory. The framework also includes the notion of a stream of observations, the modeling of the

decay of the truthfulness of individual observations, and how to integrate 'expired' and 'prevalent' observations into the agent's beliefs. More precisely, observations which are no longer immediately relevant become *default assumptions* until overridden by newer, more prevalent observations. Consider the following three scenarios.

Scenario 1 An acquaintance (Bianca) shows you the new, expensive iSung-8 smartphone she bought, with a very particular cover design. A year later, you visit Bianca at her home, where only her teenage son lives in addition. You see two phones on the kitchen counter, an iSung-8 with the particular cover design you remember from a year ago and an iSung-9. Is the likelihood greater that the iSung-8 or the iSung-9 is Bianca's?

Scenario 2 You are at the airport in a city away from home and you expect to land in your home city (Cape Town) in three hours' time. You hear someone waiting for the same flight say 'It is raining in Cape Town'. Is the likelihood less that it will still be raining if your flight is delayed by 3 hours than if the flight was not delayed?

Scenario 3 Your neighbour tells you he needs to visit the dentist urgently. You know that he uses the dentist at the Wonder-mall. A month later, you see your neighbour at the Wonder-mall. Is he there to see the dentist?

What these three scenarios have in common is that the answers to the questions make use of the persistence of truth of certain pieces of information. After a period has elapsed, the veracity of some kinds of information dissipates. For instance, in Scenario 1, one might attach an 'expiry date' to the information that the particular iSung-8 phone is Bianca's. So, by the time you visit her, the truth of that information is much weaker, in fact, it has become *defeasible* by then. Hence, we may easily argue that Bianca gave her old phone to her son and she bought the iSung-9 for herself. However, if you had visited her one month after she showed you her new iSung-8 and you saw it together with the newer iSung-9 on the counter, you would probably rather assume that the iSung-9 was Bianca's son's. In Scenario 2, you could expect it to be raining in Cape Town when you get there in three hours (because spells of rain usually last for four hours in Cape Town), but if your flight is delayed, there will be no rain or only drizzle when you land. The information 'It is raining in Cape Town' has a lifespan of four hours. With respect to Scenario 3, one would expect a person who says they must visit the dentist urgently to visit the dentist within approximately seven days. So your neighbour is probably not at Wonder-mall to see the dentist. Hence 'Neighbour

must visit dentist' should be true for no longer than seven days, after which, the statement becomes defeasibly true.

In this paper, I attempt to formalise some of these ideas. Several simplification are made; two main simplifications are (i) all information packets (evidence/observations) have a meaningful period for which they can be thought of as certainly true and (ii) the transition of a piece of information from certainly true to defeasibly true is immediate. Consider the following pieces of information.

1. Bianca is an acquaintance of mine.
2. It will rain in Cape Town this week.
3. My dentist is Dr. Oosthuizen.

The first statement is problematic because, for instance, Bianca might *gradually* become a friend. With respect to the second statement, it is easy to set the 'expiry period' to coincide with the end of the week. One might feel that it is difficult to assign a truth period to 'My dentist is Dr. Oosthuizen' due to lack of information. A person typically does not have the same dentist life-long, but one can usually not predict accurately *when* one will get a new dentist. On the other hand, if, for instance, one knows exactly when one is moving to a new city, then one can give a meaningful truth period, and the transition of the piece of information from certainly true to defeasibly true is immediate.

Many of these issues are studied in *temporal logics* [10, 11]. The focus of the present work, however, is more on belief change with a simple temporal aspect (and the integration of POMDP theory). One may do well in future research to attempt combining the results of the present work with established work on temporal logics. One paper particularly relevant to the present work presents a probabilistic temporal logic capable of modeling reasoning about evidence [7].

Expired observations are continually aggregated into the agent's set of default assumptions. Prevalent (unexpired) observations remain in the agent's 'current memory stream'. Whenever the agent wants to perform some reasoning task, it combines the prevalent observations with its fixed beliefs, then modifies its changed beliefs with respect to its default assumptions, and reasons with respect to this final set of beliefs. However, the agent always reverts back to the original fixed beliefs (hence, "fixed"). The default assumptions keep changing as memories fade, that is, as observations expire.

The rest of this paper unfolds as follows. In the next section, I review the three formalisms on which the framework is mainly based, namely, *partial probability theory*, *POMDPs* and the *hybrid stochastic belief change* framework. Section 3 presents the formal definition of the framework and Section 4 explains how the framework components interact and change when used for reasoning. A discussion about the possible patterns of cognition within the framework is presented in Section 5. Then Section 6 provides an extensive example, showing some of the computations which would be required in practice. The paper ends with some final remarks and pointers to related work.

2 FORMAL FOUNDATIONS

Let L be a finite classical propositional language. A *world* is a logical model which evaluates every propositional variable to true or false, and by extension, evaluates every propositional sentence in L to true or false. Given n propositional atoms, there are 2^n conceivable worlds. Let W be a set of possible worlds – a subset of the conceivable worlds. The fact that $w \in W$ satisfied $\psi \in L$ (or is a model for ψ) is denoted by $w \Vdash \psi$.

Let a *belief state* b be a probability distribution over all the worlds in W. That is, $b : W \to [0,1]$, such that $\sum_{w \in W} b(w) = 1$. For all $\psi \in L, b(\psi) := \sum_{w \in W, w \Vdash \psi} b(w)$.

Let L^{pc} be some *probability constraint language* which has atoms constraining the probability of some propositional sentence being true, and contains all formulae which can be formed with the atoms in combination with logical connectives. If $C \subset L^{pc}$ is a set of formulae, then b *satisfies* C (denoted $b \Vdash C$) iff $\forall \beta \in C$, b satisfies the constraints posed by β. I denote Π as the set of all belief states over the set of possible worlds W. Let Π^C be the set of belief states which satisfy the probability constraints in C. That is, $\Pi^C := \{c \in \Pi \mid c \Vdash C\}$. Later, the notion of the theory of a set of belief states will be useful:

$$Th(\Pi^C) := \{\psi \in L^{pc} \mid b \in \Pi^C, b \Vdash \psi\}.$$

I build on Voorbraak's [36] partial probability theory (PPT), which allows probability assignments to be partially determined, and where there is a distinction between probabilistic information based on (i) hard background evidence and (ii) some assumptions. An epistemic state in PPT is defined as the quadruple $\langle \Omega, \mathcal{B}, \mathcal{A}, \mathcal{C} \rangle$, where Ω is a sample space, $\mathcal{B} \subset L^{pc}$ is a set of probability constraints, $\mathcal{A} \subset L^{pc}$ is a sets of assumptions and $\mathcal{C} \subseteq W$ "represents specific information concerning the case at hand" (an observation or evidence).

Voorbraak mentions that he will only consider conditioning where the evidence does not contradict the current beliefs. He defines the set of belief states corresponding to the conditionalized PPT epistemic state as $\{b(\cdot \mid C) \in \Pi \mid b \in \Pi^{\mathcal{B} \cup \mathcal{A}}, b(C) > 0\}$.

Voorbraak proposes *constraining* as an alternative to conditioning: Let $\psi \in L^{pc}$ be a probability constraint. Then, constraining Π^B on ψ produces $\Pi^{B \cup \{\psi\}}$. Note that expanding a belief set reduces the number of models (worlds) and expanding a PPT epistemic state with extra constraints also reduces the number of models (belief states / probability functions).

> In the context of belief sets, it is possible to obtain any [... epistemic] state from the ignorant [... epistemic] state by a series of expansions. In PPT, constraining, but not conditioning, has the analogous property. This is one of the main reasons we prefer constraining and not conditioning to be the probabilistic version of expansion. [36, p. 4]

Voorbraak provides the following example [36].

Example 1. *Consider a robot which has to recharge his battery. This can be done in two rooms, let us call them room 1 and 2. The rooms are equally far away from the robot. An example of generic information might be: "the door of room 1 is at least 40% of the time open". Suppose there is no other information available, and let p_i denote the probability of door i being open. Then $\mathcal{B} = \{p_1 \geq 0.4\}$. Since doors are typically sometimes open and sometimes closed, it might be reasonable to include $0 < p_i < 1$ in $\mathcal{A}$. However, such additional assumptions should be invoked only when they are necessary, for example, in case no reasonable decision can be made without assumptions. A good example of specific evidence is information about the state of the doors obtained by the sensors of the robot.*

The word *assumption* is not very informative: An assumption may be highly entrenched (indefeasible), for instance, about the laws of physics, or an assumption may be very tentative (defeasible) like hearing gossip about some character trait of a new staff member. However, implicit in the word *default*, is the notion of defeasibility; default information is information which holds *until* stronger evidence defeats it. I shall refer to the set B as *background knowledge*

and assume it to be indefeasible, and the set F as *foreground knowledge* and assume it to be defeasible. Hence, referring to Voorbraak's Example 1, $p_1 \geq 0.4$ would be in F and $0 < p_i < 1$ would be in B. Indeed, in PPT, "it is intended to be understood that the conclusions warranted by $[\mathcal{A} \cup \mathcal{B}]$ depend on the assumptions represented in $\mathcal{A}$," [37]. I interpret this to mean that background knowledge ($\mathcal{A}$) dominates foreground knowledge ($\mathcal{B}$).

It is, however, conceivable that background knowledge should be defeasible and that new (foreground) evidence should weigh stronger due its recency and applicability. In the proposed framework, a compromise is attempted: background knowledge is 'dominated' by new evidence at the time of reasoning, but after reasoning, the new evidence is 'forgotten'. However, evidence is not completely forgotten: as it becomes less applicable after some time, it gets assimilated into the foreground knowledge as default information.

I also use elements of partially observable Markov decision process (POMDP) theory [1, 25, 23]. In a POMDP, the agent can only predict with a likelihood in which state it will end up after performing an action. And due to imperfect sensors, an agent must maintain a probability distribution over the set of possible states.

Formally, a POMDP is a tuple $\langle S, A, T, R, \Omega, O \rangle$ with a finite set of states $S = \{s_1, s_2, \ldots, s_n\}$, a finite set of actions $A = \{a_1, a_2, \ldots, a_k\}$, the *state-transition function*, where $T(s, a, s')$ is the probability of being in s' after performing action a in state s, the *reward function*, where $R(a, s)$ is the reward gained for executing a while in state s, a finite set of observations $\Omega = \{z_1, z_2, \ldots, z_m\}$; and the *observation function*, where $O(a, z, s')$ is the probability of observing z in state s' resulting from performing action a in some other state. An initial belief state b^0 over all states in S is assumed given.

To update the agent's beliefs about the world, a *state estimation function* $SE(b, a, z) = b_{a,z}^{SE}$ is defined as

$$b_{a,z}^{SE}(s') = \frac{O(a, z, s') \sum_{s \in S} T(s, a, s') b(s)}{Pr(z \mid a, b)},$$

where a is an action performed in 'current' belief-state b, z is the resultant observation and $b_{a,z}^{SE}(s')$ denotes the probability of the agent being in state s' in 'new' belief-state $b_{a,z}^{SE}$. Note that $Pr(z \mid a, b)$ is a normalizing constant.

Let the *planning horizon* h (also called the *look-ahead depth*) be the number of future steps the agent plans ahead each time it selects its next action. $V^*(b, h)$ is the *optimal* value of future courses of actions the agent can take with respect to a finite horizon h starting in belief-state b. This function assumes that at each step, the action which will maximize the state's value will be selected. $V^*(b, h)$ is defined as

$$\max_{a \in A} \left[\rho(a, b) + \gamma \sum_{z \in \Omega} Pr(z \mid a, b) V^*(SE(b, a, z), h - 1) \right],$$

where $\rho(a, b)$ is defined as $\sum_{s \in S} R(a, s) b(s)$, $0 < \gamma \leq 1$ is a factor to discount the value of future rewards and $Pr(z \mid a, b)$ denotes the probability of reaching belief-state $b_{a,z}^{SE} = SE(b, a, z)$. While V^* denotes the optimal state value, function Q^* denotes the optimal action value: $Q^*(a, b, h) = \rho(a, b) + \gamma \sum_{z \in Z} Pr(z \mid a, b) V^*(SE(b, a, z), h - 1)$ is the value of executing a in the current belief-state, plus the total expected value of belief-states reached thereafter.

I also build on my *stochastic belief change model* [29], which is a structure $\langle W, Evt, T, E, O, Str \rangle$, with a set of possible worlds W, a set of events Evt, an *event-transition function* where $T(w, e, w')$ models the probability of a transition to world w', given the occurrence of event e in world w, an event likelihood function where $E(e, w) = P(e \mid w)$ is the probability of the occurrence of event e in w, an *observation function* where $O(z, w)$ models the probability of observing z in w, and where $Str(z, w)$ is the agent's ontic strength for z perceived in w.

I proposed a way of trading off the probabilistic update and probabilistic revision, using the notion of ontic strength. The argument is that an agent could reason with a range of degrees for information being ontic (the effect of a physical action or occurrence) or epistemic (purely informative). It is assumed that the higher the informations degree of being ontic, the lower the epistemic status of that information. "An agent has a certain sense of the degree to which a piece of received information is due to a physical action or event in the world. This sense may come about due to a combination of sensor readings and reasoning. If the agent performs an action and a change in the local environment matches the expected effect of the action, it can be quite certain that the effect is ontic information," [29, p. 129].

The hybrid stochastic change of belief state b due to new information z with ontic strength (denoted $b \lhd z$) is defined as [29]

$$b \lhd z := \Big\{ (w, p) \mid w \in W, p = $$
$$\frac{1}{\gamma} \big[(1 - Str(z, w)) b_z^*(w) + Str(z, w) b_z^\diamond(w) \big] \Big\},$$

where $*$ is some probabilistic belief revision operator, $\diamond$ is some probabilistic belief update operator and γ is a normalizing factor so that $\sum_{w \in W} b_z^\lhd(w) = 1$.

The principle of maximum entropy [18, 27, 34, 26, 20] says that it is reasonable to represent a set of belief states Π^C by the member of Π^C which is most entropic or least biased (w.r.t. information theory) of all the members of Π^C:

$$ME(\Pi^C) := \arg\max_{c \in \Pi^C} H(c),$$

where $H(c) := -\sum_{w \in W} c(w) \ln c(w)$.

The principle of minimum cross-entropy [22, 5] is used to select the belief state $c \in \Pi^Y$ 'most similar' to a given belief state $b \in \Pi^X$; the principle minimizes the *directed divergence* between b and c with respect to entropy. Directed divergence is defined as

$$R(c, b) := \sum_{w \in W} c(w) \ln \frac{c(w)}{b(w)}.$$

$R(c, b)$ is undefined when $b(w) = 0$ while $c(w) > 0$; when $c(w) = 0$, $R(c, b) = 0$, because $\lim_{x \to 0} \ln(x) = 0$.

The knowledge and reasoning structure proposed in this paper is presented next. It builds on the work of Voorbraak [36] and Rens [29] and includes elements of POMDP theory.

3 FRAMEWORK DEFINITION

Let L^{prob} be a probabilistic language over L defined as $L^{prob} := \{\phi[\ell, u] \mid \phi \in L, \ \ell, u \in [0, 1], \ell \leq u\}$. A sentence of the form $\phi[\ell, u]$ means the likelihood of proposition ϕ is greater than or equal to ℓ and less than or equal to u. Let $\mathbb{N} = \{0, 1, 2, \ldots\}$.

b *satisfies* formula $\phi[\ell, u]$ (denoted $b \Vdash \phi[\ell, u]$) iff $\ell \leq b(\phi) \leq u$.

Definition 1. *An agent maintains a structure $\langle W, B, F, A, Evt, Z, Prs, Eng, \mathcal{M} \rangle$, where*

- W *is a set of possible worlds;*

- $B \subset L^{prob}$ *is a* background belief base *of fixed assumptions;*
- $F \subset L^{prob}$ *is a* foreground belief base *of default assumptions;*
- A *is a set of (agent)* actions*, including a special action* null*;*
- Evt *is a set of (environment)* events*;*
- Z *is the* observation stream*, a set of observation triples:* $Z := \{(a_1, t_1, z_1), (a_2, t_2, z_2), \ldots, (a_k, t_k, z_k)\}$*, where* $a_i \in A$*,* $t_i \in \mathbb{N}$*,* $z_i \in L$*, and such that* $\forall t_i, t_j \in \mathbb{N}$*,* $i = j$ *iff* $t_i = t_j$ *(i.e., no more than one action and observation occur at a time-point);*
- $Prs : L \times W \to \mathbb{N}$ *is a* persistence function*, where* $Prs(z, w)$ *indicates how long* z *is expected to be true from the time it is received, given the 'context' of* w*; it is a total function over* $L \times W$*;*
- $Eng : L \times W \times A \to [0, 1]$*, where* $Eng(z, w, a)$ *is the agent's confidence that* z *perceived in* w *was caused by action* a *(i.e., that* z *has an endogenous source);*
- $\mathcal{M}$ *is a model of the environment, and any auxiliary information required by the definition of the particular belief change operation* $(\lhd)$*.*

Definition 2. *The expected persistence of* z *perceived in belief state* b *is*

$$ExpPrs(z, b) := \sum_{w \in W} Prs(z, w) \times b(w).$$

But the agent will reason with respect to a *set* of belief states Π^C, hence, $ExpPrs(z, \Pi^C)$ must be defined. One such definition employs the principle of maximum entropy:

Definition 3.

$$ExpPrs^{ME}(z, \Pi^C) := \sum_{w \in W} Prs(z, w) \times b^{ME}(w),$$

where $b^{ME} = ME(\Pi^C)$*.*

Definition 4. *An observation triple* $(a, i, z) \in Z$*, has* expired *at point* s *if* $ExpPrs(z, b) < s - i$*.*

Let $b_{a,z}^{\lhd}$ be the *change* of belief state b by a and z. My intention is that "change" is a neutral term, not necessarily indicating revision or update. Next, I propose one instantiation of $b_{a,z}^{\lhd}$. Let the environment model $\mathcal{M} = \langle E, T, O \rangle$, where

- $E : Evt \times W \to [0, 1]$ is the *event function*. $E(e, w) = P(e \mid w)$, the probability of the occurrence of event e in w;
- $T : W \times (A \cup Evt) \times W \to [0, 1]$ is a *transition function* such that for every $\ae \in A \cup Evt$ and $w \in W$, $\sum_{w' \in W} T(w, \ae, w') = 1$, where $T(w, \ae, w')$ models the probability of a transition to world w', given the execution of action / occurrence of event $\ae$ in world w;
- $O : W \times W \times A \to [0, 1]$ is an *observation function* such that for every $w \in W$ and $a \in A$, $\sum_{w^z \in W} O(w^z, w, a) = 1$, where $O(w^z, w, a)$ models the probability of observing ϕ^z (a complete theory for w^z) in w and where $O(z, w, a) := \sum_{\substack{w^z \in W \\ w^z \Vdash z}} O(w^z, w, a)$, for all $z \in L$.

Observation z may be due to an *exogenous* event (originating and produced outside the agent) or an *endogenous* action (originating and produced within the agent). It is up to the agent designer to decide, for each observation, whether it is exogenous or endogenous, given the action and world.

The hybrid stochastic belief change (HSBC) formalism of Rens

[29] defines the (exogenous) update of b with z as

$$b_z^\diamond := \left\{ (w, p) \mid w \in W, p = \frac{1}{\gamma} O(z, w) \sum_{w' \in W} \sum_{e \in Evt} E(e, w') T(w', e, w) b(w') \right\},$$

where γ is a normalizing factor and $O(z, w)$ can be interpreted as $O(z, w, null)$. But HSBC does not involve agent actions; HSBC assumes that agents passively receive information. Hence, when an agent is assumed to act and the actions are known, the POMDP state estimation function can be employed for belief state update.

Only for the propose of illustrating how the framework can be used, the following belief change procedure is defined.

$$b_{a,z}^{\lhd} := \{ (w, p) \mid w \in W, p = Exg(z, w, a) b_z^\diamond(w) + Eng(z, w, a) b_{a,z}^{SE}(w) \}, \quad (1)$$

where $Exg(z, w, a) := 1 - Eng(z, w, a)$ is the confidence that z is exogenous in w, given a was executed.

4 OPERATIONAL SEMANTICS

The agent is always living at time-point $\mathcal{N}$. Time units remain the same and must be chosen to suit the domain of interest, for instance, milliseconds, minutes, hours, etc. The first point in time is $\mathcal{N} = 0$. $\mathcal{N}$ indicated the end of the $\mathcal{N}$-th time unit, in other words, at point $\mathcal{N}$, exactly $\mathcal{N} \times u$ time has passed, where u is the time unit employed.

It will be assumed that no observation can be made at time-point 0. If the agent designer feels that it is unreasonable for the agent to have to wait until $\mathcal{N} = 1$ before the first observation may be made, then the time unit chosen for the particular domain is too large.

Expired observations are continually aggregated into the agent's set of default assumptions F (foreground beliefs). Prevalent (unexpired) observations remain in the agent's 'current memory stream' Z. Whenever the agent wants to perform some reasoning task, (i) it combines the prevalent observations with its fixed beliefs B, then modifies its changed beliefs with respect to its default assumptions, or (ii) modifies its fixed beliefs B with respect to its default assumptions, then combines the prevalent observations with its changed beliefs – and then reasons with respect to this final set of beliefs. However, the agent always reverts back to the original fixed beliefs (hence, "fixed"). And the default assumptions keep changing as memories fade, that is, as observations expire.

Because any initial conditions specified are, generally, not expected to hold after the initial time-point, it does not make sense to place them in B. The only other option is to place them in F, which is allowed to change. The following guiding principle is thus provided to agent designers.

The agent's initial beliefs (i.e., the system conditions at point $\mathcal{N} = 0$) must be specified in F.

Let $C \subset L^{prob}$ be a belief base. At this early stage of research, I shall suggest only three definitions of $\lhd$ on a set of belief states. The first is the most naïve approach (denoted $\lhd_{NV}$). It is suitable for theoretical investigations.

$$\Pi^C \lhd_{NV} a, z := \{ b_{a,z}^{\lhd} \mid b \in \Pi^C \}.$$

A more practical approach is to reduce Π^C to a representative belief state, employing the principle of maximum entropy (denoted

$\lhd_{ME}$).

$$\Pi^C \; \lhd_{ME} \; a, z := \{b_{a,z}^{\lhd} \mid b = ME(\Pi^C)\}.$$

A third and final approach which I shall mention here is, in a sense, a compromise between the naïve and maximum entropy approaches. It is actually a family of methods which will thus not be defined precisely. It is the approach (denoted $\lhd_{FS}$) which finds a finite (preferably, relatively small) proper subset Π^{FS} of Π^C which is somehow representative of Π^C, and then applies $\lhd$ to the individual belief states in Π^{FS}:

$$\Pi^C \; \lhd_{FS} \; a, z := \{b_{a,z}^{\lhd} \mid b \in \Pi^{FS} \subset \Pi^C\}.$$

Only $\lhd_{ME}$ and $\lhd_{FS}$ will be considered in the sequel, when it comes to belief change over a set. Let $\lhd_{set}$ denote one of these two operators. Then $\Pi^C \; \lhd_{set} \; \sigma$ can be defined, where σ is any stream of observation triples, where $\sigma = \{(a_1, t_1, z_1), (a_2, t_2, z_2), \ldots, (a_k, t_k, z_k)\}$ and $t_1 < t_2 < \cdots < t_k$:

$$\Pi^C \leftarrow \Pi^C \; \lhd_{set} \; a_1, z_1$$
$$\text{and then } \Pi^C \leftarrow \Pi^C \; \lhd_{set} \; a_2, z_2 \ldots$$
$$\text{and then } \Pi^C \leftarrow \Pi^C \; \lhd_{set} \; a_k, z_k.$$

Let *Expired* be derived from the triples in Z which have just expired (at point $\mathcal{N}$). That is,

$$Expired := \{(a, i, z) \in Z \mid ExpPrs(z, \Pi_{\mathcal{N}-1}) < \mathcal{N} - i\},$$

where $\Pi_{\mathcal{N}}$ is defined below. At each time-point, F is refreshed with all the expired observation triples in the order they appear in Z. In other words, at each point in time, $F \leftarrow Th(\Pi^F \; \lhd_{set} \; Expired)$. As soon as the foreground has been refreshed, the expired triples are removed from the observation stream: $Z \leftarrow Z \setminus Expired$. I shall use the notation $\Pi^F_{\lhd_{set}}$ to clarify which operation was used to arrive at the current Π^F.

A function which selects a belief state in one set which is in some sense closest to another set of belief states will shortly be required. The following definition suggests three such functions.

Definition 5.

$$Closest^{absolute}(\Pi^X, \Pi^Y) := \underset{b: \, b \in \Pi^X, c \in \Pi^Y}{\arg\min} \sum_{w \in W} (b(w) - c(w))^2$$

or

$$Closest^{ME}(\Pi^X, \Pi^Y) := \underset{b \in \Pi^X}{\arg\min} \sum_{w \in W} (b(w) - ME(\Pi^Y)(w))^2,$$

or

$$Closest^{MCE}(\Pi^X, \Pi^Y) := \underset{b: \, b \in \Pi^X, c \in \Pi^Y}{\arg\min} R(c, b),$$

where $X, Y \in L^{prob}$, $ME(\Pi^Y)$ is the most entropic/least biased belief state in Π^Y and $R(c, b)$ is the directed divergence of c with respect to b.

$Closest^{absolute}$ simply chooses the belief state $b \in \Pi^X$ which minimizes the sum of the differences between probabilities of worlds of b and some belief state in $c \in \Pi^Y$ (considering all $c \in \Pi^Y$).

$Closest^{ME}$ picks the belief state $c^{ME} \in \Pi^Y$ with maximum entropy (i.e., least biased w.r.t. information in Π^Y), and then chooses the belief state $b \in \Pi^X$ which minimizes the sum of the differences between probabilities of worlds of b and c^{ME}.

$Closest^{MCE}$ chooses the belief state $b \in \Pi^X$ for which the directed divergence from belief state $c \in \Pi^Y$ with respect to entropy is least (considering all $c \in \Pi^Y$). This is an instance of the minimum cross-entropy inference.

From here onwards, I shall not analyse their definitions; $Closest(\cdot)$ will be used to refer to the abstract function.

Reasoning is done with respect to the set of belief states $\Pi_{\mathcal{N}}$. I look at two patterns of cognition to determine $\Pi_{\mathcal{N}}$.

Definition 6.

$$(\Pi_{\mathcal{N}})_1 := \begin{cases} (\Pi^B \lhd_{set} Z) \cap \Pi^F_{\lhd_{set}} \; if \, (\Pi^B \lhd_{set} Z) \cap \Pi^F_{\lhd_{set}} \neq \emptyset \\ \{Closest(\Pi^B \lhd_{set} Z, \Pi^F_{\lhd_{set}})\} \; otherwise, \end{cases}$$

or

$$(\Pi_{\mathcal{N}})_2 := \begin{cases} (\Pi^B \cap \Pi^F_{\lhd_{set}}) \lhd_{set} Z \; if \, \Pi^B \cap \Pi^F_{\lhd_{set}} \neq \emptyset \\ \{\{Closest(\Pi^B, \Pi^F_{\lhd_{set}})\} \lhd_{set} Z\} \; otherwise. \end{cases}$$

The idea behind the definition of $(\Pi_{\mathcal{N}})_1$ is that (pertinent) observations in the stream (Z) dominate background beliefs (B), but Z-modified beliefs ($\Pi^B \lhd_{set} Z$) dominate foreground beliefs (F). The idea behind the definition of $(\Pi_{\mathcal{N}})_2$ is that foreground beliefs (F) have slightly higher status than in $(\Pi_{\mathcal{N}})_1$ because they modify background beliefs (when consistent with B) *before* (pertinent) observations in the stream (Z), but the stream finally dominates.

5 PATTERNS OF COGNITION

In this section, I shall explore the properties of 'patterns of cognition' based on $(\Pi_{\mathcal{N}})_1$ and $(\Pi_{\mathcal{N}})_2$. I shall argue that there are four reasonable candidates. First, a few axioms:

1. $\{b\} \cap \Pi^C = \{b\}$ or $\emptyset$.
2. $\{b\} \cap \{c\} \neq \emptyset \iff b = c$.
3. $\Pi^C \lhd_{ME} Z$ always results in a singleton set.
4. $Closest(\{b\}, \Pi^C) = b$.

And one guiding principle:

In a given pattern of cognition, $\lhd_{set}$ is instantiated to the same operator when it appears in the same relative position.

For instance, in $(\Pi_{\mathcal{N}})_1$,

$$(\Pi^B \lhd_{ME} Z) \cap \Pi^F_{\lhd_{FS}} \; if \, (\Pi^B \lhd_{ME} Z) \cap \Pi^F_{\lhd_{FS}} \neq \emptyset$$
$$\{Closest(\Pi^B \lhd_{ME} Z, \Pi^F_{\lhd_{FS}})\} \; otherwise,$$

is allowed, but not

$$(\Pi^B \lhd_{ME} Z) \cap \Pi^F_{\lhd_{FS}} \; if \, (\Pi^B \lhd_{ME} Z) \cap \Pi^F_{\lhd_{FS}} \neq \emptyset$$
$$\{Closest(\Pi^B \lhd_{FS} Z, \Pi^F_{\lhd_{FS}})\} \; otherwise,$$

Or in $(\Pi_{\mathcal{N}})_2$,

$$(\Pi^B \cap \Pi^F_{\lhd_{FS}}) \lhd_{FS} Z \; if \, \Pi^B \cap \Pi^F_{\lhd_{FS}} \neq \emptyset$$
$$\{Closest(\Pi^B, \Pi^F_{\lhd_{FS}})\} \lhd_{FS} Z \; otherwise.$$

is allowed, but not

$$(\Pi^B \cap \Pi^F_{\lhd_{ME}}) \lhd_{FS} Z \; if \, \Pi^B \cap \Pi^F_{\lhd_{FS}} \neq \emptyset$$
$$\{Closest(\Pi^B, \Pi^F_{\lhd_{FS}})\} \lhd_{FS} Z \; otherwise.$$

Given these axioms and the guiding principle, four reasonable candidates are

$$(\Pi_{\mathcal{N}})_{11} := \begin{cases} (\Pi^B \lhd_{FS} Z) \cap \Pi^F_{\lhd_{FS}} \text{ if } (\Pi^B \lhd_{FS} Z) \cap \Pi^F_{\lhd_{FS}} \neq \emptyset \\ \{Closest(\Pi^B \lhd_{FS} Z, \Pi^F_{\lhd_{FS}})\} \text{ otherwise,} \end{cases}$$

$$(\Pi_{\mathcal{N}})_{12} := \{Closest(\Pi^B \lhd_{FS} Z, \Pi^F_{\lhd_{set}})\}.$$

$$(\Pi_{\mathcal{N}})_{21} := \begin{cases} (\Pi^B \cap \Pi^F_{\lhd_{FS}}) \lhd_{set} Z \text{ if } \Pi^B \cap \Pi^F_{\lhd_{FS}} \neq \emptyset \\ \{Closest(\Pi^B, \Pi^F_{\lhd_{FS}})\} \lhd_{set} Z \text{ otherwise.} \end{cases}$$

$$(\Pi_{\mathcal{N}})_{22} := \{Closest(\Pi^B, \Pi^F_{\lhd_{set}}) \lhd Z\}.$$

Note that in the definitions of $(\Pi_{\mathcal{N}})_{12}$, $(\Pi_{\mathcal{N}})_{21}$ and $(\Pi_{\mathcal{N}})_{22}$, $\lhd_{set}$ may be instantiated as either of the two operators, and in $(\Pi_{\mathcal{N}})_{22}$, $\lhd$ is actually the 'plain' belief change operator.

I now justify each of the four patterns of cognition.

11 If $(\Pi^B \lhd_{FS} Z)$ were $(\Pi^B \lhd_{ME} Z)$, by axiom 3, the result is a singleton set, and by axioms 1 and 4, the information in Π^F is ignored, including initial conditions or their later effects. If $\Pi^F_{\lhd_{FS}}$ were $\Pi^F_{\lhd_{ME}}$, by axiom 1, the information in $(\Pi^B \lhd_{set} Z)$ is ignored, or worse, the intersection is empty.

12 With this pattern, the issues of intersection (axioms 1 and 2) need not be dealt with. If $\Pi^B \lhd_{FS} Z$ were $\Pi^B \lhd_{ME} Z$, by axiom 3, the result is a singleton set, and by axiom 4, the information in Π^F is ignored. Whether $\Pi^F_{\lhd_{set}}$ is instantiated as $\Pi^F_{\lhd_{ME}}$ or $\Pi^F_{\lhd_{FS}}$, the set still has an influence on $\Pi^B \lhd_{FS} Z$ due to the latter typically not being a singleton.

21 If $(\Pi^B \cap \Pi^F_{\lhd_{FS}})$ were $(\Pi^B \cap \Pi^F_{\lhd_{ME}})$, by axioms 1, the information in Π^B is ignored. Whether $\lhd_{set} Z$ is instantiated as $\lhd_{ME} Z$ or $\lhd_{FS} Z$, the result will typically not be trivial – all information is taken into account, although in different ways.

22 Either of the two instantiations of $\Pi^F_{\lhd_{set}}$ accommodate the information in Π^B and in Π^F. Moreover, the issues of intersection (axioms 1 and 2) need not be dealt with, and due to the simpler pattern, the second belief change operation can also be simplified, that is, in this pattern, $\lhd_{ME} Z$ and $\lhd_{FS} Z$ both reduce to $\lhd Z$.

6 AN EXAMPLE

To keep things as simple as possible for this introductory report, I shall illustrate $(\Pi_{\mathcal{N}})_{22}$ instantiated as

$$Closest^{absolute}(\Pi^B, \Pi^F_{\lhd_{ME}}) \lhd Z.$$

Consider the following scenario. The elves live in the forest. Sometimes, orcs (monsters) wander into the forest to hunt or seek shelter from a storm. Elves don't like orcs. The elves can either do nothing (action null) or they can evict the orcs (action evict). Orcs tend to stay in the forest for more hours the darker it is. The sun rises (event rise) and sets (event set) once a day (duh). Now and then, there is a storm (event storm), but most of the time, nothing happens (event null). The vocabulary will be $\{e, o, \ell\}$ meaning, respectively, the elves are in the forest, the orcs are in the forest, and it is light (it is daytime and there isn't a storm).

Henceforth, I might write $\alpha\beta$ instead of $\alpha \wedge \beta$, and $\bar{\alpha}$ instead of $\neg\alpha$. For ease of reading, the possible worlds will be ordered, and a belief state $\{(eo\ell, p_1), (eo\bar{\ell}, p_2), (e\bar{o}\ell, p_3), (e\bar{o}\bar{\ell}, p_4), (\bar{e}o\ell, p_5), (\bar{e}o\bar{\ell}, p_6), (\bar{e}\bar{o}\ell, p_7), (\bar{e}\bar{o}\bar{\ell}, p_8)\}$ will be abbreviated as $\langle p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8\rangle$.

Only observations $eo, e\bar{o}, \ell, \bar{\ell}$ are considered.

Let

- $Eng(z, w, \text{null}) = 0, \forall z \in \Omega, \forall w \in W$ (Observations after null actions are never endogenous).
- $Eng(eo, w, \text{evict}) = 0, \forall w \in W$ (Orcs in the forest should never be observed *due to* eviction).
- $Eng(e\bar{o}, w, \text{evict}) = 0.75$ if $w \Vdash \bar{\ell}$, = 0.5 if $w \Vdash \ell$ (One can be more confident that the orcs are not in the forest due to the eviction, when it is dark).
- $Eng(\ell, w, \text{evict}) = 0, \forall w \in W$ (Lightness is independent of eviction).

I am not entirely comfortable with this model of endogeny; see the concluding section for a short discussion.

Let

- $Prs(eo) = \{(\ell, 2), (\bar{\ell}, 4)\}$ (Once the orcs are observed in the forest, one can rely on them remaining there for at least two hours when light and four hours when dark).
- $Prs(e\bar{o}) = \{(\top, 1)\}$ (One can rely on the orcs remaining outside the forest for one hour once one finds out that they are outside).
- $Prs(\ell) = \{(\top, 5)\}$ (One can rely on light for five hours once light is perceived; a storm may darken things within five ours).
- $Prs(\bar{\ell}) = \{(\top, 3)\}$ (There might be a daytime storm, which may clear up within three hours from when it stared and was perceived).

To define the belief change operator $\lhd$ as in (1), $\mathcal{M} = \langle E, T, O\rangle$ must be defined as follows.

- If $w \Vdash \ell$,
  - $E(\text{rise}, w) = 0$,
  - $E(\text{set}, w) = 1/12$,
  - $E(\text{storm}, w) = 3/12$.
  - $E(\text{null}, w) = 8/12$.
- If $w \Vdash \neg\ell$,
  - $E(\text{rise}, w) = 1/12$,
  - $E(\text{set}, w) = 0$,
  - $E(\text{storm}, w) = 3/12$.
  - $E(\text{null}, w) = 8/12$.
- $T(w, \text{null}, w) = 1 \; \forall w \in W$ (for the action and event).
- $T(w, \text{evict}, w') = 0$ if $w \Vdash \neg e \vee \neg o$, else if $w \Vdash \ell$,
  - $= 0.1$ if $w' \Vdash e \wedge o \wedge \ell$,
  - $= 0.9$ if $w' \Vdash e \wedge \neg o \wedge \ell$,
  else if $w \Vdash \neg\ell$,
  - $= 0.1$ if $w' \Vdash e \wedge o \wedge \neg\ell$,
  - $= 0.9$ if $w' \Vdash e \wedge \neg o \wedge \neg\ell$.
- $T(w, \text{rise}, w') = 1$ if $w \Vdash \neg\ell$ and $w' \Vdash \ell$ and truth values of e and o are invariant.
- $T(w, \text{set}, w') = 1$ if $w \Vdash \ell$ and $w' \Vdash \neg\ell$ and truth values of e and o are invariant.
- $T(w, \text{storm}, w') = 1$ if $w' \Vdash \neg\ell$ and truth values of e and o are invariant.

The probabilities for $O(z, w, \text{null})$ are as follows.

z \ w	$eo\ell$	$eo\bar{\ell}$	$e\bar{o}\ell$	$e\bar{o}\bar{\ell}$	$\bar{e}o\ell$	$\bar{e}o\bar{\ell}$	$\bar{e}\bar{o}\ell$	$\bar{e}\bar{o}\bar{\ell}$
eo	0.8	0.8	0.1	0.1	0.1	0.1	0	0
$e\bar{o}$	0.1	0.1	0.8	0.8	0	0	0.1	0.1
ℓ	1	0	1	0	1	0	1	0
$\bar{\ell}$	0	1	0	1	0	1	0	1

The probabilities for $O(z, w, \texttt{evict})$ are as follows.

z \ w	$eo\ell$	$eo\bar{\ell}$	$e\bar{o}\ell$	$e\bar{o}\bar{\ell}$	$\bar{e}o\ell$	$\bar{e}o\bar{\ell}$	$\bar{e}\bar{o}\ell$	$\bar{e}\bar{o}\bar{\ell}$
eo	0.8	0.8	0	0	0	0	0	0
$e\bar{o}$	0.2	0.2	0.8	0.8	0	0	0.2	0.2
ℓ	1	0	1	0	1	0	1	0
$\bar{\ell}$	0	1	0	1	0	1	0	1

Initially, the foreground belief base will specify the initial condition,

$$F = \{\ell[1,1], eo[.7, .9]\},$$

that is, it is completely light and there is a $70\% - 90\%$ belief that the elves and orcs are in the forest. And the background belief base demands that the probability that elves are outside the forest while orcs are inside is never more than 10%,

$$B = \{\bar{e}o[0, 0.1]\}.$$

Let us start by analyzing the effect of observation stream

$$Z = \{(\texttt{null}, 1, eo), (\texttt{evict}, 4, e\bar{o})\}$$

at different time-points.

At $\mathcal{N} = 1$, $(\Pi_1)_{22} = Closest^{absolute}(\Pi^B, \Pi^F_{\lhd_{ME}}) \lhd \{(\texttt{null}, 1, eo)\}$ (given $Prs(eo) = \{(\ell, 2), (\bar{\ell}, 4)\}$, $(\texttt{null}, 1, eo)$ has not yet expired). $\Pi^F_{\lhd_{ME}} = ME(\Pi^F) = \langle .7, 0, .1, 0, .1, 0, .1, 0 \rangle$. Due to Π^B and $\Pi^F_{\lhd_{ME}}$ being mutually consistent, $Closest^{absolute}(\Pi^B, \Pi^F_{\lhd_{ME}})$ is simply $\langle .7, 0, .1, 0, .1, 0, .1, 0 \rangle \in \Pi^B$. Denote $\langle .7, 0, .1, 0, .1, 0, .1, 0 \rangle$ as b_1.

$Eng(eo, w, \texttt{null}) = 0$ for all $w \in W$. Therefore, $b_1 \lhd \{(\texttt{null}, 1, eo)\} = b_1 \lhd \texttt{null}, eo = (b_1)^\diamond_{eo}$, which was calculated to be $\langle .386, .579, .007, .01, .007, .01, 0, 0 \rangle$ (denoted b_2 henceforth). I shall only show how $(b_1)^\diamond_{eo}(eo\bar{\ell}) = 0.579$ is calculated. Note that, due to the transition functions for the four events being invariant with respect to e and o, the only departure worlds we need to consider are (abusing notion) $eo\ell$ and $eo\bar{\ell}$: $(b_1)^\diamond_{eo}(eo\bar{\ell}) =$

$$\frac{1}{\gamma} O(eo, eo\bar{\ell}, \texttt{null}) \sum_{w' \in W} \sum_{e \in Evt} b_1(w') E(e, w') T(w', e, eo\bar{\ell})$$

$$= \frac{1}{\gamma}(0.8) \sum_{w' \in W} b_1(w') \times [E(\texttt{null}, w') T(w', \texttt{null}, eo\bar{\ell})$$

$$+ E(\texttt{rise}, w') T(w', \texttt{rise}, eo\bar{\ell})$$

$$+ E(\texttt{set}, w') T(w', \texttt{set}, eo\bar{\ell})$$

$$+ E(\texttt{storm}, w') T(w', \texttt{storm}, eo\bar{\ell})]$$

$$= \frac{1}{\gamma}(0.8) \big[b_1(eo\ell) [E(\texttt{null}, eo\ell) T(eo\ell, \texttt{null}, eo\bar{\ell})$$

$$+ E(\texttt{rise}, eo\ell) T(eo\ell, \texttt{rise}, eo\bar{\ell})$$

$$+ E(\texttt{set}, eo\ell) T(eo\ell, \texttt{set}, eo\bar{\ell})$$

$$+ E(\texttt{storm}, eo\ell) T(eo\ell, \texttt{storm}, eo\bar{\ell})]$$

$$+ b_1(eo\bar{\ell}) [E(\texttt{null}, eo\bar{\ell}) T(eo\bar{\ell}, \texttt{null}, eo\bar{\ell})$$

$$+ E(\texttt{rise}, eo\bar{\ell}) T(eo\bar{\ell}, \texttt{rise}, eo\bar{\ell})$$

$$+ E(\texttt{set}, eo\bar{\ell}) T(eo\bar{\ell}, \texttt{set}, eo\bar{\ell})$$

$$+ E(\texttt{storm}, eo\bar{\ell}) T(eo\bar{\ell}, \texttt{storm}, eo\bar{\ell})]\big]$$

$$= \frac{1}{\gamma}(0.8) \big[(0.7)[8/12 + 0 + 1/12 + 3/12]$$

$$+ (0)[8/12 + 0 + 0 + 3/12]\big]$$

$$= \frac{1}{\gamma} 0.56.$$

It turns out that $\gamma = 0.968$, resulting in $0.56/0.968 = 0.579$.

The agent believes to a high degree what it perceived (eo). The reason why it believes to a higher degree that it is dark than light, is due to the relatively high chance of a storm (which darkens things) and a small chance of the sun setting. This scenario is a bit synthetic: the agent has not yet perceived that it is light. In a realistic situation, the agent will always sense the brightness of the environment, disallowing a high degree of belief in darkness.

At $\mathcal{N} = 5$, $Expired = \{(\texttt{null}, 1, eo)\}$ because $ExpPrs(eo, b_1) = \sum_{w \in W} Prs(eo, w) \times b_1(w) = 3.2 < \mathcal{N} - 1 = 4$. And $(\texttt{evict}, 4, e\bar{o}) \notin Expired$ because $ExpPrs(e\bar{o}, b_1) = \sum_{w \in W} Prs(e\bar{o}, w) \times b_1(w) = 1 \not< \mathcal{N} - 4 = 1$.

Therefore, $(\Pi_5)_{22} = Closest^{absolute}(\Pi^B, \Pi^F_{\lhd_{ME}}) \lhd \{(\texttt{evict}, 4, e\bar{o})\} = Closest^{absolute}(\Pi^B, \{b_1 \lhd \texttt{null}, eo\}) \lhd \{(\texttt{evict}, 4, e\bar{o})\} = Closest^{absolute}(\Pi^B, \{b_2\}) \lhd \{(\texttt{evict}, 4, e\bar{o})\}$.

$b_2 \in \Pi^B$, hence, $Closest^{absolute}(\Pi^B, \{b_2\}) = b_2$. $(\Pi_5)_{22}$ is thus $b_2 \lhd \texttt{evict}, e\bar{o}$.

Recall that $Eng(e\bar{o}, w, \texttt{evict})$ equals 0.75 if $w \Vdash \bar{\ell}$, but 0.5 if $w \Vdash \ell$. And recall that

$$b^\lhd_{a,z}(w) = Exg(z, w, a) b^\diamond_z(w) + Eng(z, w, a) b^{SE}_{a,z}(w).$$

So, for instance,

$$(b_2)^\lhd_{\texttt{evict}, e\bar{o}}(eo\bar{\ell})$$

$$= Exg(e\bar{o}, eo\bar{\ell}, \texttt{evict})(b_2)^\diamond_{e\bar{o}}(eo\bar{\ell})$$

$$+ Eng(e\bar{o}, eo\bar{\ell}, \texttt{evict}) b^{SE}_{\texttt{evict}, e\bar{o}}(eo\bar{\ell})$$

$$= 0.25(b_2)^\diamond_{e\bar{o}}(eo\bar{\ell}) + 0.75(b_2)^{SE}_{\texttt{evict}, e\bar{o}}(eo\bar{\ell})$$

At $\mathcal{N} = 6$, $Expired = \{(\texttt{null}, 1, eo), (\texttt{evict}, 4, e\bar{o})\}$ because $(\texttt{null}, 1, eo)$ had already expired at $\mathcal{N} = 5$, and $ExpPrs(e\bar{o}, (b_2)^\lhd_{\texttt{evict}, e\bar{o}}) = 1 < \mathcal{N} - 4 = 2$.

Therefore, $(\Pi_6)_{22} = Closest^{absolute}(\Pi^B, \Pi^F_{\lhd_{ME}}) \lhd \{\} = Closest^{absolute}(\Pi^B, \{(b_2)^\lhd_{\texttt{evict}, e\bar{o}}\})$.

7 CONCLUDING REMARKS

I believe that it is important to have the facility to reason about both (exogenous) events and (endogenous) actions; I am definitely not the first to propose a framework with both notions [28, 33, 9, 6, e.g.]. The framework also has two belief bases, one to represent fixed, background beliefs and one to accommodate defeasible information (observations which have become 'stale', but not necessarily false).

Inherent to the framework is that the agent's knowledge may be incomplete. There is much work on dealing with ignorance or missing information [14, 16, 37, 38, 21, 30, e.g.].

What makes the proposed framework potentially significant is its generality, applicable to many domains and agent-designer requirements. I want to amplify the point that the belief change operations used in this paper are only suggestions and used due to personal familiarity with them – the researcher / agent designer is given the flexibility to suggest or design their own operators to suit their needs.

Another feature of this framework which potentially adds to its significance, is the nature of the observation stream (with expiring observations) and how it interacts with the dual belief base approach. We saw the rich potential of patterns of cognition, which can be simultaneously confusing and empowering to the framework user. However, some of the confusion was cleared up in Section 5. My feeling is that this observation-stream-dual-belief-base system holds

much potential for investigating deep questions in how to model continual observation and cognition within a belief change setting.

One line of research that should be made is to generalize the dual-belief-base approach: What would happen if three, four, more belief bases are employed, each accommodating a different degree of entrenchment of given and received information? Would such a generalization reduce to the known systems of belief change (which include notions of preference, plausibility, entrenchment, etc.)?

Closely related to the discussion in the previous paragraph is that keeping the belief base B fixed is quite a strong stance. In reality, only the most stubborn people will never change their core views even a little bit. Such stubbornness indicates an inability to grow, that is, an inability to improve one's reasoning and behaviour. In the current framework, the plasticity of the assumptions, although important for accommodating and aggregating recent observations, are always dominated by B. In future versions, I would like to make B more amenable to learning, while minding sound principles of belief change in logic and cognitive psychology.

Perhaps one of the most difficult aspects of using the proposed framework is to specify the persistence function $Prs(\cdot)$. However, the specification is made easier by the property of the operational semantics that: expired observations keep on having an influence of the agent's reasoning until (if) they are 'overridden' by the process of refreshing set F. This means that the agent designer should rather err by specifying less persistence of observations when s/he is uncertain about the period to specify. In other words, the agent designer is advised to specify the longest period an observation is *guaranteed* to persist.

I also found it challenging to thinking about how to model how endogenous `evict` is, for the different worlds and observations. For instance, if $Eng(e\bar{o}, w, \texttt{evict}) = 0.75$ when $w \Vdash \bar{\ell}$, should it constrain the values that $Eng(\bar{\ell}, w, \texttt{evict})$ can take? And if it is impossible for the elves to evict while they are outside the forest, can $Eng(z, w, \texttt{evict})$ be greater than 0 if $w \Vdash \neg e$? There seem to be several inter-related issues in the modeling of endogeny, including action executability, perceivability and consistency among related observations. I did not focus on these issues here.

It would be straightforward to add a utility function (e.g., a POMDP reward function) to the environment model $\mathcal{M}$. Existing planning and decision-making algorithms and methods can then be used together with the proposed framework. Instead of using the POMDP state estimation function during POMDP planning, for instance, a more general belief change operator ($\lhd$) could be used in planning under uncertainty, where the operator's definition depends on the elements of the proposed framework. Little work on planning and decision-making with underspecified knowledge exists [31], [36, Sec. 5].

The fundamental distinction between *focusing* and *belief revision* when dealing with generic knowledge has been made by Dubois and Prade [8]: "Revision amounts to modifying the generic knowledge when receiving new pieces of generic knowledge (or the factual evidence when obtaining more factual information), while focusing is just applying the generic knowledge to the reference class of situations which exactly corresponds to all the available evidence gathered on the case under consideration." The distinction seems very relevant to general systems for knowledge management in dynamic environments.

This paper touches on several aspects of computational reasoning, including stochasticity, imprecision/ignorance, knowledge entrenchment, default knowledge, physical change and belief update, new evidence and belief revision, and the persistence of evidence.

Except for evidence persistence, there are probably hundreds of papers and article on combinations of these aspects. I could not yet find any work dealing with the persistence of veracity of new evidence/observations, as presented in the present paper. Besides the work already cited in this paper, the following may be used as a bibliography to better place the present work in context, and to point to methods, approaches and techniques not covered in the proposed framework, which could possibly be added to it.

- Probabilistic logics for reasoning with defaults and for belief change or learning [15, 24].
- Nonmonotonic reasoning systems with optimum entropy inference as central concept [4, 2, 3].
- Dynamic epistemic logics for reasoning about probabilities [35, 32].

ACKNOWLEDGEMENTS

REFERENCES

[1] K. Aström, 'Optimal control of Markov decision processes with incomplete state estimation', *Journal of Mathematical Analysis and Applications*, **10**, 174–205, (1965).

[2] C. Beierle and G. Kern-Isberner, 'On the modelling of an agent's epistemic state and its dynamic changes', *Electronic Communications of the European Association of Software Science and Technology*, **12**, (2008).

[3] C. Beierle and G. Kern-Isberner, 'Towards an agent model for belief management', in *Advances in Multiagent Systems, Robotics and Cybernetics: Theory and Practice. (Volume III)*, eds., G. Lasker and J. Pfalzgraf, IIAS, Tecumseh, Canada, (2009).

[4] R. Bourne and S. Parsons, 'Extending the maximum entropy approach to variable strength defaults', *Ann. Math. Artif. Intell.*, **39**(1–2), 123–146, (2003).

[5] I. Csiszár, 'I-divergence geometry of probability distributions and minimization problems', *Annals of Probability*, **3**, 146–158, (1975).

[6] F. Dupin de Saint-Cyr and J. Lang, 'Belief extrapolation (or how to reason about observations and unpredicted change)', *Artif. Intell.*, **175**, (2011).

[7] D. Doder, Z. Marković, Z. Ognjanović, A. Perović, and M. Rašković, *A Probabilistic Temporal Logic That Can Model Reasoning about Evidence*, 9–24, Lecture Notes in Computer Science, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[8] D. Dubois and H. Prade, *Focusing vs. belief revision: A fundamental distinction when dealing with generic knowledge*, 96–107, Springer Berlin Heidelberg, Berlin, Heidelberg, 1997.

[9] A. Ferrein, C. Fritz, and G. Lakemeyer, 'On-line decision-theoretic Golog for unpredictable domains', in *KI 2004: Advances in Artif. Intell.*, volume 238/2004, 322–336, Springer Verlag, Berlin / Heidelberg, (2004).

[10] D. Gabbay, I. Hodkinson, and M. Reynolds, *Temporal Logic: Mathematical Foundations and Computational Aspects*, volume 1, Clarendon Press, Oxford, 1994.

[11] D. Gabbay, M. Reynolds, and M. Finger, *Temporal Logic: Mathematical Foundations and Computational Aspects*, volume 2, Oxford University Press, Oxford, 2000.

[12] P. Gärdenfors, *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, MIT Press, Massachusetts/England, 1988.

[13] P. Gärdenfors, *Belief Revision*, volume 29 of *Cambridge Tracts in Theoretical Computer Science*, Cambridge University Press, Massachusetts/England, 1992.

[14] H. Geffner and B. Bonet, 'High-level planning and control with incomplete information using POMDPs', in *Proceedings of the Fall AAAI Symposium on Cognitive Robotics*, pp. 113–120, Seattle, WA, (1998). AAAI Press.

[15] M. Goldszmidt and J. Pearl, 'Qualitative probabilities for default reasoning, belief revision, and causal modeling', *Artificial Intelligence*, **84**, 57–112, (1996).

[16] A. Grove and J. Halpern, 'Updating sets of probabilities', in *Proceedings of the Fourteenth Conf. on Uncertainty in Artif. Intell.*, UAI'98, pp. 173–182, San Francisco, CA, USA, (1998). Morgan Kaufmann.

[17] S. Hansson, *A textbook of belief dynamics: theory change and database updating*, Kluwer Academic, Dortrecht, The Netherlands, 1999.

[18] E. Jaynes, 'Where do we stand on maximum entropy?', in *The Maximum Entropy Formalism*, 15–118, MIT Press, (1978).

[19] H. Katsuno and A. Mendelzon, 'On the difference between updating a knowledge base and revising it', in *Proceedings of the Second Intl. Conf. on Principles of Knowledge Representation and Reasoning*, pp. 387–394, (1991).

[20] G. Kern-Isberner, 'Characterizing the principle of minimum cross-entropy within a conditional-logical framework', *Artif. Intell.*, **98**(12), 169 – 208, (1998).

[21] G. Kern-Isberner, 'Linking iterated belief change operations to nonmonotonic reasoning', in *Proceedings of the Eleventh Intl. Conf. on Principles of Knowledge Representation and Reasoning*, pp. 166–176, Menlo Park, CA, (2008). AAAI Press.

[22] S. Kullback, *Information theory and statistics*, volume 1, Dover, New York, 2nd edn., 1968.

[23] W. Lovejoy, 'A survey of algorithmic methods for partially observed Markov decision processes', *Annals of Operations Research*, **28**, 47–66, (1991).

[24] T. Lukasiewicz, 'Nonmonotonic probabilistic logics under variable-strength inheritance with overriding: Complexity, algorithms, and implementation', *International Journal of Approximate Reasoning*, **44**(3), 301–321, (2007).

[25] G. Monahan, 'A survey of partially observable Markov decision processes: Theory, models, and algorithms', *Management Science*, **28**(1), 1–16, (1982).

[26] J. Paris, *The Uncertain Reasoner's Companion: A Mathematical Perspective*, Cambridge University Press, Cambridge, 1994.

[27] J. Paris and A. Vencovsk, 'In defense of the maximum entropy inference process', *Intl. Journal of Approximate Reasoning*, **17**(1), 77–103, (1997).

[28] D. Poole, 'Decision theory, the situation calculus and conditional plans', *Linköping Electronic Articles in Computer and Information Science*, **8**(3), (1998).

[29] G. Rens, 'On stochastic belief revision and update and their combination', in *Proceedings of the Sixteenth Intl. Workshop on Non-Monotonic Reasoning (NMR)*, eds., G. Kern-Isberner and R. Wassermann, pp. 123–132. Technical University of Dortmund, (2016).

[30] G. Rens, T. Meyer, and G. Casini, 'Revising incompletely specified convex probabilistic belief bases', in *Proceedings of the Sixteenth Intl. Workshop on Non-Monotonic Reasoning (NMR)*, eds., G. Kern-Isberner and R. Wassermann, pp. 133–142. Technical University of Dortmund, (2016).

[31] G. Rens, T. Meyer, and G. Lakemeyer, 'A modal logic for the decision-theoretic projection problem', in *Proceedings of the Seventh Intl. Conf. on Agents and Artif. Intell. (ICAART), Revised Selected Papers*, eds., B. Duval, J. Van den Herik, S. Loiseau, and J. Filipe, LNAI, pp. 3–19. Springer Verlaag, (2015).

[32] J. Sack, 'Extending probabilistic dynamic epistemic logic', *Synthese*, **169**, 124–257, (2009).

[33] S. Shapiro and M. Pagnucco, 'Iterated belief change and exogenous actions in the situation calculus', in *Proceedings of the Sixteenth European Conf. on Artif. Intell. (ECAI-04)*, pp. 878–882, Amsterdam, (2004). IOS Press.

[34] J. Shore and R. Johnson, 'Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy', *Information Theory, IEEE Transactions on*, **26**(1), 26–37, (Jan 1980).

[35] J. Van Benthem, J. Gerbrandy, and B. Kooi, 'Dynamic update with probabilities', *Studia Logica*, **93**(1), 67–96, (2009).

[36] F. Voorbraak, 'Partial Probability: Theory and Applications', in *Proceedings of the First Intl. Symposium on Imprecise Probabilities and Their Applications*, pp. 360–368, (1999).

[37] F. Voorbraak, 'Probabilistic belief change: Expansion, conditioning and constraining', in *Proceedings of the Fifteenth Conf. on Uncertainty in Artif. Intell.*, UAI'99, pp. 655–662, San Francisco, CA, USA, (1999). Morgan Kaufmann Publishers Inc.

[38] A. Yue and W. Liu, 'Revising imprecise probabilistic beliefs in the framework of probabilistic logic programming.', in *Proceedings of the Twenty-third AAAI Conf. on Artif. Intell. (AAAI-08)*, pp. 590–596, (2008).