Interpreting an action from what we perceive and what we expect.

January 16, 2006

Abstract

In update logic as studied by Baltag, Moss, Solecki and van Benthem, little attention is paid to the interpretation of an action, which is just assumed to depend on the situation. This is actually a complex issue that nevertheless complies to some logical dynamics. In this paper, we tackle this topic. We also deal with actions that change propositional facts of the situation. In Parallel, we propose a formalism to accurrately represent an agent's epistemic state based on hyperreal numbers. In that respect, we use infinitesimals to express what would surprise the agents (and by how much) by contradicting their beliefs. We also use a subjective probability to model the notion of belief. It turns out that our probabilistic update mechanism satisfies the AGM postulates of belief revision.

The interpretation of an action is a complex process that is rather neglected in the literature, in which it is at most assumed to depend on the situation. To get a glimpse of the logical dynamics underlying this process, let us have a look at two examples. Assume that you see somebody drawing a ball from an urn containing black balls and white balls. If you *believe* there is no particular distribution in the urn then you would expect with equal probability that he draws a white ball or a black ball; but if you *believe* there are more black than white balls in the urn then you would expect with a higher probability that he draws a black ball rather than a white ball. We see in this example that your beliefs of the situation contribute actively to interpret the action: they determine the probability with which you would expect (or would have expected) the action to happen. But this expectation, determined by your beliefs of the situation, can often be balanced consciously or unconsciously by what you actually obtain by the pure observation of the action. For example, assume that you listen to a message of one of your colleagues on your answering machine which says that he will come to your office on Tuesday afternoon, but you cannot distinguish precisely due to some noise whether he said Tuesday or Thursday. From your beliefs about his schedule on Tuesday and Thursday, you would have expected him to say that he would come on Tuesday (if you believe he is busy on Thursday) or on Thursday (if you believe he is busy on Tuesday). But this expectation has to be balanced by what you actually perceive and distinguish from the message on the answering machine: you might also have a preference between having heard Tuesday or Thursday, which is independent of this expectation. Thus we see in these examples that in the process of interpreting an action, there is an interplay between two main informational components: your expectation of the action to happen and your pure perception of the action happening. Up to now, this kind of phenomenon, although very common in everyday life but also essential when we want realistic updates, is not dealt with neither in update logic ([BMS04], [vBen03], [Auc05a]) nor in the situation calculus ([BaHaLe98]), nor in a previous version of this paper ([Auc05b]). For sake of generality, the actions we consider in this paper can also change (propositional) facts of the situation.

In order to represent the agents' epistemic state accurately we introduce a formalism based on hyperreal numbers defined in Sect. 1. This enables us to model both degrees of belief and degrees of potential surprise, thanks to infinitesimals. The richness of this formalism will then allow us to tackle the dynamics of belief appropriately, notably revision, and also to express in a suitable language what would surprise the agents (and by how much) by contradicting their beliefs.

In this paper we follow the approach of update logic as viewed by Baltag, Moss and Solecki (BMS) ([BMS04]). So as in BMS, we divide our task in three parts. Firstly, we propose a formalism called pd-model to represent how the world is perceived by the agents from a static point of view (Sect. 2.1). Secondly, we propose a formalism called generic action model to represent how an action occurring in this world is perceived by the agents (Sect. 2.2). Thirdly, we propose an update mechanism which takes as arguments a pd-model and a generic action model, and yields a new pd-model; the latter represents how the world is perceived by the agents after the action represented by the above generic action model took place in the world represented by the above pd-model (Sect. 2.3). Finally we give some comparisons with the AGM postulates and relevant literature (Sect. 3).

1 Mathematical Preliminaries

In our system, the probabilities of worlds and formulas will take value in a particular mathematical structure $(\mathbb{V}, \lessapprox)$ (abusively noted $(\mathbb{V}, \leq)$) different from the real numbers, based on hyperreal numbers $({}^*\mathbb{R}, \leq)$ ([Adams75] uses them as well to give a probabilistic semantics to conditional logic). First, we will briefly recall the main features of hyperreal numbers that will be useful in the sequel; for details see [Keis86]. Afterwards we will motivate and introduce our particular structure $(\mathbb{V}, \lessapprox)$.

Roughly speaking, hyperreal numbers are an extension of the real numbers to include certain classes of infinite and infinitesimal numbers. A hyperreal number x is said to be infinitesimal iff $|x| < 1/n$ for all integers n, finite iff $|x| < n$ for some integer n, and infinite iff $|x| > n$ for all integers n. Infinitesimal numbers are typically denoted ε, finite numbers are denoted x and infinite numbers are denoted ∞. Note that an infinitesimal number is a finite number as well. Note also that $\frac{1}{\varepsilon}$ is an infinite number and $\frac{1}{\infty}$ is an infinitesimal number.

Two hyperreal numbers b and c are said to be infinitely close to each other if their difference $b - c$ is infinitesimal. If x is finite, the standard part of x, denoted by $St(x)$, is the unique real number which is infinitely close to x. So for example $St(1 + \varepsilon) = 1$, $St(\varepsilon) = 0$.

What we would like to do is to approximate our expressions. That is to say, in case a hyperreal number a is infinitely smaller than b, i.e. there is an infinitesimal ε such that $a = \varepsilon.b$, then we would want $b + a = b$. For example we would want $1 + \varepsilon = 1$ (here $a = \varepsilon$ and $b = 1$), $\varepsilon + \varepsilon^2 = \varepsilon$ (here $a = \varepsilon^2$ and $b = \varepsilon$),... In other words, in case a is *negligible* compared to b, then $b + a = b$. Unfortunately, the hyperreal numbers do not allow us to do that, so we are obliged to devise a new structure $(\mathbb{V}, \lessapprox)$.

First we introduce some definitions for a better clarity. By *semi-field* (resp. *ordered semi-field*), we mean a field (resp. ordered field) which lacks the property of 'existence of additive inverse'. For example, $(\mathbb{R}^+, +, ., \leq)$ and $({}^*\mathbb{R}^+, +, ., \leq)$ are ordered semi-fields, where ${}^*\mathbb{R}^+$ denotes the positive hyperreal numbers. Now we define $\mathbb{V}$, which will be the quotient structure of the set of positive hyperreal numbers ${}^*\mathbb{R}^+$ by a particular equivalence relation.

Definition 1.1 Let $x, y \in {}^*\mathbb{R}^+$, we set

$$x \approx y \text{ iff } \begin{cases} St(\frac{x}{y}) = 1 & \text{if } y \neq 0 \\ x = 0 & \text{if } y = 0. \end{cases}$$

We can easily check that $\approx$ is an equivalence relation on ${}^*\mathbb{R}^+$.

$\lhd$

Theorem 1.2 *The quotient structure* $\mathbb{V} = ({}^*\mathbb{R}^+/_\approx, \overline{+}, \overline{.})$ *is a semi-field.*

The proof of this theorem and of all the other lemmas and theorems in this paper are in the appendix. Now we need to define the ordering relation $\lessapprox$ on $\mathbb{V}$.

Definition 1.3 We define a relation $\lessapprox$ on $\mathbb{V}$ by
$\overline{x} \lessapprox \overline{y}$ iff there are $x \in \overline{x}$, $y \in \overline{y}$ such that $x \leq y$ $\lhd$

$\mathbb{V}$ equipped with $\lessapprox$ turns out to be an ordered semi field thanks to the following lemma:

Lemma 1.4 *If* $\overline{x} \lessapprox \overline{y}$ *then for all* $x' \in \overline{x}$ *and all* $y' \in \overline{y}$, $x' \leq y'$

Theorem 1.5 *The structure* $(\mathbb{V}, \lessapprox)$ *is an ordered semi-field.*

We will denote abusively $\overline{x}$, $\overline{+}$, $\overline{.}$, $\lessapprox$ by $x, +, ., \leq$. The elements $\overline{x}$ containing an infinitesimal will be called abusively *infinitesimals* and denoted $\varepsilon, \delta, \ldots$. Those containing a real number will be called abusively *reals* and denoted $a, b, \ldots$ Those containing an infinite number will be called *infinites* and denoted $\infty, \infty', \ldots$ Moreover, when we refer to intervals, these intervals will be in $\mathbb{V}$; so for example $]0; 1]$ refers to $\{a \in \mathbb{V}; 0 \lessapprox a \lessapprox 1$ and not $a \approx 0\}$.

We can then easily check that now we do have at our disposal the following identities: $1 + \varepsilon = 1, a + \varepsilon = a, \varepsilon + \varepsilon^2 = \varepsilon, \ldots$

It turns out that our structure $\mathbb{V}$ is an extension of a structure $\mathbb{V}'$ isomorphic to a cumulative algebra, which is a notion introduced by Weydert ([Wey94]).

Theorem 1.6 $\mathbb{V}' = \{x \in \mathbb{V}; x$ *is real or* x *is infinitesimal* $\}$ *is isomorphic to a cumulative algebra.*

2 Dynamic Proba-doxastic Logic

2.1 The Static Part

2.1.1 The Notion of pd-model.

Definition 2.1 A proba-doxastic model (pd-model) is a tuple $M = (W, \{\sim_j; j \in G\}, \{P_j; j \in G\}, V, w_0)$ where:

1. W is a finite set of possible worlds;

2. w_0 is the possible world corresponding to the actual world;

3. $\sim_j$ is an equivalence relation defined on W for each agent j;

4. P_j is a probabilistic measure defined for each agent j on each set $[w]_j := \{v; v \sim_j w\}$ which assigns to each world $v \in [w]_j$ a number in $]0; 1]$ and such that $\sum\{P_j(v); v \sim_j w\} = 1$ (*);

5. V is a valuation;

6. G is a set of agents.

$\lhd$

Intuitive Interpretations. Items 1,2,5,6 are clear. It remains to give interpretations for items 3 and 4. The probabilistic operator P_j together with $\sim_j$ intuitively model the epistemic state of any agent $j \in G$. We could replace $\sim_j$ and P_j by a single equivalent object P_j^w (see remark 2.2) but we consider that the following intuitive interpretations are clearer with these two components.

We set $w \sim_j v$ iff agent j cannot distinguish world w from world v. Note that this relation does not model the notion of knowledge and we do not deal with this notion in this paper.

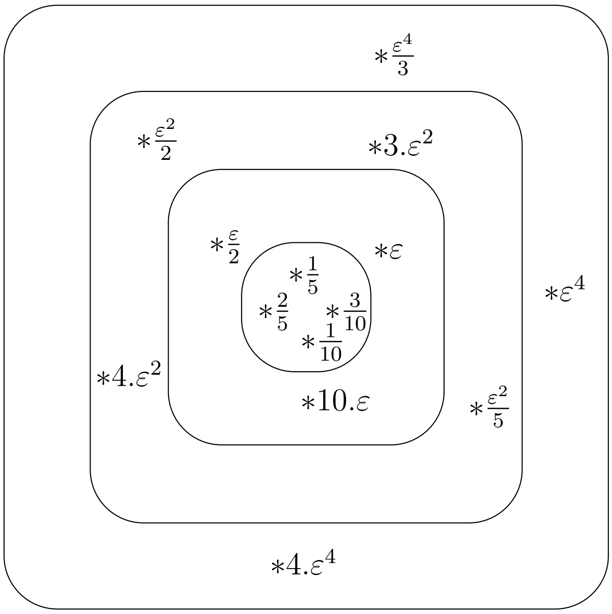
Among the worlds that agent j can not distinguish (that is to say the worlds $[w]_j$), there are some that j conceives as potential candidates for the world in which j dwells, and some that j would be surprised to hear somebody claiming that they correspond to the world in which j dwells (whatever it is). The first ones are called *conceived worlds* and the second *surprising worlds*. The conceived worlds are assigned by P_j a real value and the surprising worlds are assigned an infinitesimal value (both different from 0). For example, some people would be surprised to hear somebody claiming that some swans are black, although it is true. So for them the actual world is a surprising world, indistinguishable from the only conceived world where all swans are white.

Of course j might conceive some (conceived) worlds as better candidates than others and this is expressed by the value of the probability of the world: the larger the real probability value of the (conceived) world is, the more likely it is for j. But that is the same for the surprising worlds: j might be more surprised to hear some worlds than others. For example, if you play poker with somebody you trust, you will never suspect that he cheats. However he does so, and so carefully that you do not suspect anything. Then at the end of the game if he announces to you that he has cheated, you will be surprised (although it is something true in the actual world, which is then a surprising world). But you will be even more surprised if he tells you that he has cheated five times than if he has cheated only once. So the world where he has cheated five times will be more surprising than the world where he has cheated once, and these are both surprising worlds for you. Infinitesimals enable us to express this: the larger the infinitesimal probability value of the (surprising) world is, the less j would be surprised by this world. Anyway, that is why we need to introduce hyperreal numbers: to express these degrees of potential surprise that can not be expressed by a single number like 0 (which then becomes useless for us).

Finally, the natural condition (*) ensures us that $([w]_j, \mathcal{P}([w]_j), P_j)$ is a (nonstandard) probability space. Moreover, since P_j takes its values in $\mathbb{V}'$ which is isomorphic to a cumulative algebra (see Sect. 1), P_j is a cumulative measure in the sense of Weydert ([Wey94]).

So, dwelling in one of the worlds of $[w]_j$, j does not think consciously that her respective surprising worlds in $[w]_j$ are possible (unlike conceived worlds), she is just not aware of them. So they are useless to represent her beliefs which we assume are essentially conscious. But still, these worlds are relevant for the (external and objective) modelling. Indeed they provide some information about the epistemic state of mind of j : namely what would surprise her and how firmly she holds to her belief. Intuitively, something that you do not consider consciously as possible and that contradicts your beliefs is often surprising for you when it is claimed by somebody else. These worlds will moreover turn out to be very useful technically in case j has to revise her beliefs (see Sect. 3).

In the picture below is depicted an example of a $\sim_j$ equivalence class. The asterisks correspond to worlds and the numbers next to them represent their respective probabilities. The conceived worlds are the ones in the inner circle, the other worlds are surprising worlds. Note that the sum of the probabilities of the conceived worlds is equal to 1. But the global sum of all these worlds (conceived and surprising) is also equal to 1 because the sum of a real number with an infinitesimal is equal to this real number (see Sect.1). The presence of the other circles will be explained in Sect.3, they are related to Weydert's notion of ranking.



Remark 2.2 Let (W, P_j^w, V, w_0) be a model where for each world w and each agent j, $P_j^w : W \longrightarrow [0;1]$ is a probabilistic operator such that for all $w, v \in W$

1. $P_j^w(w) > 0$
2. if $P_j^w(v) > 0$ then for all $u \in W$ $P_j^w(u) = P_j^v(u)$
3. $\sum\{P_j^w(v); v \in W\} = 1$.

Then we can show that every model of this kind is equivalent to a pd-model, and vice versa, that every pd-model is equivalent to a model of this kind (Lang, private communication).

2.1.2 Static Language and Examples.

We can define naturally a language $\mathcal{L}_{St}$ for pd-models (St standing for Static).

Definition 2.3 The syntax of the language $\mathcal{L}_{St}$ is defined by
$\varphi := \perp \mid p \mid \neg\varphi \mid \varphi \wedge \psi \mid P_j(\varphi) \geq x \mid C_j\varphi$ where $x \in [0;1[$.
Its semantics is defined by,
$M, w \models P_j(\varphi) \geq x$ iff $\sum\{P_j(v); w \sim_j v \text{ and } M, v \models \varphi\} \geq x$
$M, w \models C_j\varphi$ iff $\sum\{P_j(v); w \sim_j v \text{ and } M, v \models \varphi\} = 1$. $\lhd$

$M, w \models P_j(\varphi) \geq x$ should be read "in world w and for j, φ has a probability greater than x". $M, w \models C_j\varphi$ should be read "in world w, j is convinced (sure) of φ". We use the term conviction instead of belief because the term "belief" refers in natural language to different concepts that we distinguish here through $\mathcal{L}_{St}$. For example, assume that you conjecture an arithmetical theorem φ from a series of examples and particular cases. The more examples you have checked, the more you will "believe" in this theorem. This notion of belief corresponds to the type of formula $P_j(\varphi) = x$ (or $P_j(\varphi) \geq x$) for x real smaller than 1; the bigger x is the more you "believe" in φ. But if you come up with a proof of this theorem that you have checked several times, you will still "believe" in this theorem but this time with a different strength. Your belief will be similar to a conviction and corresponds to the type of formula $C_j\varphi$. However, note that this conviction (belief) might still be false if there is a mistake in the proof that you did not notice (like for the Fermat's theorem). For a more in depth account on the different significations of the term "belief", see for example [Len78]. Nevertheless, in the sequel we will alternatively and equivalently employ the terms "belief" and "conviction" for better readability. This notion of conviction actually corresponds to the Lenzen's in [Len78]. Moreover, if we define the operator B_j by $M, w \models B_j\varphi$ iff $M, w \models P_j(\varphi) > 0.5$, then B_j corresponds to the Lenzen's notion of "weak belief" and C_j and B_j satisfy the Lenzen's axioms defined in [Len78]. This notion of conviction also corresponds to Gardenfors' notion of accepted formula (in a belief set) ([Gard88]). This operator C_j satisfies the axioms K, D, 4, 5 but not the axiom T. So it does not correspond to the notion of knowledge and we recall that we do not deal with this notion in this paper.

Moreover, we can also express in this language what would surprise the agent, and by how much. Indeed, in case x is an infinitesimal ε, $M, w \models (P_j(\varphi) = \varepsilon)$ should be read "in world w, j would be surprised with intensity ε to hear somebody claiming that φ "[1]. (Note that the smaller x is, the higher the intensity of surprise is.) But this use of infinitesimals could also express how firmly we believe something, in the spirit of Spohn. Indeed, $P_j(\neg\varphi) = \varepsilon > \varepsilon' = P_j(\neg\varphi')$ would then mean that φ' is believed more firmly than φ.

Note that above, if x is real then only the conceived worlds have to be considered in the sum of the expression $\sum\{P_j(v); w \sim_j v \text{ and } M, v \models \varphi\} \geq x$ because the sum of any real (different from 0) with an infinitesimal is equal to this real (see Sect. 1). Likewise, the semantics of C_j amounts to say that φ is true only in all the *conceived* worlds of $[w]_j$ for the same reason. So it is quite possible to have a surprising actual world where $\neg\varphi$ is true and still j being convinced of φ (i.e. $C_j\varphi$): just take φ:="All swans are white" for example.

In this paper we will follow step by step two examples: the urn example and the answering machine example of the introduction.

[1]Note that this does not necessarily model *all* the things that would surprise the agent since there might still be things that the agent might conceive as possible with a low probability and still be surprised to hear them claimed by somebody else (Gerbrandy, private communication).

$$\boxed{\boxed{w_0 : p_0}} \overset{Y}{\text{———}} \boxed{w_1 : p_1} \overset{Y}{- - -} \boxed{w_{n-1} : p_{n-1}} \overset{Y}{\text{———}} \boxed{w_n : p_n}$$

$$P_Y(x) = \frac{1}{n+1} \text{ for all worlds } x$$

Figure 1: 'urn' Example

$$\boxed{\boxed{w : p}} \overset{Y}{\text{———}} \boxed{v : \neg p}$$

$$P_Y(w) = \frac{4}{5} \text{ and } P_Y(v) = \frac{1}{5}$$

Figure 2: Answering machine Example

Example 2.4 [urn example 1] Suppose we are in a fair, and there is an urn containing $n = 2.k > 0$ balls which are either white or black. You do not know how many black balls there are in the urn but you believe there is no particular distribution in the urn. Now say there is actually 0 black ball in the urn. This situation is depicted in Fig.1, where p_i stands for: "there are i black balls in the urn" and the double bordered world is the actual world. All the lines between worlds (which are not all depicted) represent equivalence relations $\sim_Y$, with Y standing for You.

Example 2.5 [answering machine example] Assume you are a teacher at university and you come back after lunch to your office. You ate with a colleague who just told you at lunch about his new schedule for this year. However you do not remember quite precisely what he said, in particular you are a bit uncertain whether his 1.5 hour lecture at 2.00 pm is on Tuesday or on Thursday, and you believe with probability $\frac{4}{5}$ (resp. $\frac{1}{5}$) that his lecture is on Tuesday (resp. Thursday). In fact it is on Tuesday at 2.00 pm. We model this situation in the model of Fig.2., where p (resp. $\neg p$) stands for "your colleague has his 1.5 hour lecture on Tuesday (resp. Thursday)"[2] and the double bordered world is the actual world.

2.2 The Dynamic Part

Definition 2.6 A generic action model is a structure $\mathbf{\Sigma} = (\Sigma, S, \{\sim_j; j \in G\}, P_j, \{P_j^\Gamma; \Gamma$ is a maximal consistent subset of S and $j \in G\}, \{Post_\sigma; \sigma \in \Sigma\}, \sigma_0)$ where

1. Σ is a finite set of possible actions;

2. σ_0 is the actual action;

[2]Note that strictly speaking we would need two propositional variables p and p', where p (resp. p') would stand for "your colleague has his lecture on Tuesday (resp.Thursday)" because the fact that he does not have lecture on Tuesday ($\neg p$) should not imply logically that he has lecture on Thursday (p'). We introduce only one propositional variable only to simplify the notations, without losing any generality.

3. $\sim_j$ is an equivalence relation defined on Σ for each agent j;

4. S is a set of formulas of $\mathcal{L}_{St}$ and their negations;

5. P_j^Γ is a probabilistic measure indexed by each agent j and each maximal consistent subset Γ of S, defined on each equivalence class $[\sigma]_j = \{\tau; \sigma \sim_j \tau\}$ and assigning to each possible action σ a *real* number in $[0;1]$, such that for each possible action σ and agent j_0

   if $P_{j_0}^\Gamma(\sigma) = 0$ then $P_j^\Gamma(\sigma) = 0$ for all $j \in G$ (**);

6. P_j is a probabilistic measure indexed by each agent j defined on each equivalence class $[\sigma]_j = \{\tau; \sigma \sim_j \tau\}$ and assigning to each possible action σ a number in $]0;1]$;

7. $Post_\sigma$ is a function which takes as argument any propositional letter p and yields a formula of $\mathcal{L}_{St}$;

8. G is a set of agents.

$\lhd$

Intuitive Interpretation. Items 1,2,3,8 are similar to definition 1. It remains to give interpretation to items 4,5,6,7 of the definition.

Item 4 and 5. S corresponds to the set of formulas (and their negation) that are relevant for the agents in a particular world where Γ is true in order to determine the probabilities P_j^Γ of Item 5.

$P_j^\Gamma(\sigma)$ is the probability that j *would (have) expect(ed)* σ to happen (among the possible actions of $[\sigma]_j$), if j assumed that she was in a world where the formulas of Γ are true. In other words, $P_j^\Gamma(\sigma)$ can thus be viewed as the probability for j that the action σ would occur (among the possible actions of $[\sigma]_j$) in a world where the formulas of Γ are true. Note that this is a conditional (probability) of the form $P_j(\sigma|\Gamma)$. Moreover, because we assume the agents to be rational, the determination of the value of this probability can often be done objectively and coincides with j's subjective determination (see examples).

This probability value is real and not infinitesimal (unlike P_j), and we can have $P_j^\Gamma(\sigma) = 0$. This last case intuitively means that the action σ can not *physically* be performed in a world where Γ is true. This condition of impossibility is inherent to the action itself and is also common knowledge among the agents, that is why we have condition (**). These operators P_j^Γ generalize the binary and rigid notion of precondition in [BMS04], [Auc05a] and [vBen03].

Instead of referring directly to worlds w of a pd-model, we refer to maximal consistent subsets Γ of a set S. This is for two reasons. First, the determination of the probability $P_j^w(\sigma)$ does not depend on all the information provided by w but often on just a part of it expressed by Γ. Second, we do not want our action models to depend on a particular pd-model and we want to be able to iterate the action; the use of maximal consistent sets allows us to do so. So from now on, we note $\boldsymbol{P_j^w(\sigma) := P_j^\Gamma(\sigma)}$ **for the unique $\boldsymbol{\Gamma}$ such that $\boldsymbol{M, w \models \Gamma}$.**

Item 6. $P_j(\sigma)$ is the probability for j that σ actually occurs among $[\sigma]_j$, determined solely by j's *perception* and *observation* of the action. The determination of this probability is based on a pure observation and perception of the action, without taking into account j's beliefs of the static situation which could alter and modify this determination consciously or unconsciously. This probability is thus independent of the probability that j would have expected σ to happen, because to determine this last probability j takes into account what she believes about the static situation (this fact will be relevant in Sect.2.3.).

Just as in the static case, we have *conceived* actions (which are assigned a real number) and *surprising* actions (which are assigned an infinitesimal). The former are possible actions that the agent conceives as possible candidates while one of the indistinguishable actions actually takes place. The latter are possible actions that j would be surprised to hear somebody claiming that they took place while one of the indistinguishable actions actually took place. For example, if you play poker with somebody you trust and at a certain point he cheats while you do not suspect anything, then the actual action of cheating will be a surprising action for you (of value ε) and will be indistinguishable for you from the conceived action where nothing particular happens (of value 1). Just as in the static case, the relative strength for j of the indistinguishable actions (conceived and surprising) is expressed by the value of the operator P_j.

Item 7. The function $Post_\sigma$ deals with the problem of determining what facts will be true in a world after the action σ takes place. Intuitively, $Post_\sigma(p)$ represents the necessary and sufficient precondition in any world w for p to be true after the performance of σ in this world w.

Example 2.7 [urn example] Consider the action whereby I draw a ball from the urn (which is actually white) and put it in my pocket, you see me doing that but you can not see the ball. This action is depicted in Fig.3. The maximal consistent sets are represented by their 'positive' components, so $\{p_i\}$ refers to the set $\{p_i, \neg p_k; k \neq i\}$.

Action σ (resp. τ) stands for "I draw a black (resp. white) ball and put it in my pocket". Clearly you can not distinguish σ from τ. Moreover, the observation and perception of the action in itself does not provide you any reason to have a preference between me drawing a black ball or a white ball; so we set $P_Y(\sigma) = P_Y(\tau) = \frac{1}{2}$.[3]

[3]We could nevertheless imagine some exotic situation where your observation of me drawing a ball would give you some information on the color of the ball I am drawing. For example, if black balls were much heavier than white balls and you see me having difficulty drawing a ball, you could consider

$$\boxed{\sigma} \overset{Y}{\longrightarrow\!\!\!\!\!-} \boxed{\boxed{\tau}}$$

$$P_Y(\sigma) = P_Y(\tau) = \tfrac{1}{2}.$$
$$S = \{p_i, \neg p_i; i = 0..n\}$$
$$P_Y^{\{p_i\}}(\sigma) = \tfrac{i}{n},\ P_Y^{\{p_i\}}(\tau) = 1 - \tfrac{i}{n} \text{ for all } i.$$
$$Post_\sigma(p_n) = \perp \text{ and } Post_\sigma(p_i) = p_{i+1} \text{ if } i < n$$
$$Post_\tau(p_n) = \perp \text{ and } Post_\tau(p_i) = p_i \text{ if } i < n$$

Figure 3: I draw a (white) ball and put it in my pocket and you are uncertain whether I draw a white or a black ball.

However if you assumed you were in a world where there are i black balls then the probability that you would (have) expect(ed) me drawing a black (resp. white) ball would be $\frac{i}{n}$ (resp. $1 - \frac{i}{n}$); so we set $P_Y^{\{p_i\}}(\sigma) = \frac{i}{n}$ and $P_Y^{\{p_i\}}(\tau) = 1 - \frac{i}{n}$. Moreover there can not be n black balls in the urn after I put one ball in my pocket; so we set $Post_\sigma(p_n) = \perp$ and $Post_\tau(p_n) = \perp$. But if I draw a black ball then there is one black ball less; so we set $Post_\sigma(p_i) = p_{i+1}$ for all $i < n$. Otherwise if I draw a white ball the number of black balls remains the same; so we set $Post_\tau(p_i) = p_i$ for all $i < n$.

Example 2.8 [answering machine] Assume now that when you enter in your office, you find a message on your answering machine from your colleague. He tells you that he will bring you back a book he had borrowed you next Tuesday in the beginning of the afternoon between 2.00 pm and 4.00 pm. However, there is some noise on the message and you can not distinguish precisely whether he said Tuesday or Thursday. Nevertheless you have a slight preference for having heard Tuesday rather than Thursday.

We model this action in Fig.4. σ stands for "your colleague says that he will come on Tuesday between 2.00 pm and 4.00 pm" and τ stands for "your colleague says that he will come on Thursday between 2.00 pm and 4.00 pm". $P_Y(\sigma)$ and $P_Y(\tau)$ represent the probabilities you assign to σ and τ on the sole basis of what you have heard and recognized from the answering machine. These probabilities are then determined on the basis of your sole perception of the message. On the other hand, $P_Y^{\{p\}}(\sigma)$ is the probability that you would have expected your colleague to say that he will come on Tuesday (rather than Thursday) if you assumed that his lecture was on Tuesday. This probability can be determined objectively. Indeed, because we assume that your colleague has a lecture on Tuesday from 2.00 pm to 3.30 pm, the only time he could come on Tuesday would be between 3.30 pm and 4.00 pm (only 0.5 hour). So we set $P_Y^{\{p\}}(\sigma) = \frac{0.5hr}{2.5hrs} = \frac{1}{5}$. Similarly, $P_Y^{\{p\}}(\tau) = \frac{2hrs}{2.5hrs} = \frac{4}{5}$. Finally, the message does not change the fact of him having a lecture or not on Tuesday; so we set $Post_\sigma(p) = p$.

more probable that I am drawing a black ball.

$$\boxed{\boxed{\sigma}} \overset{Y}{\text{———}} \boxed{\tau}$$

$$P_Y(\sigma) = \tfrac{3}{5}, P_Y(\tau) = \tfrac{2}{5}$$
$$S = \{p, \neg p\}.$$
$$P_Y^{\{p\}}(\sigma) = \tfrac{1}{5},\ P_Y^{\{p\}}(\tau) = \tfrac{4}{5}$$
$$P_Y^{\{\neg p\}}(\sigma) = \tfrac{4}{5},\ P_Y^{\{\neg p\}}(\tau) = \tfrac{1}{5}.$$
$$Post_\sigma(p) = p, \text{ and } Post_\tau(p) = p$$

Figure 4: You are uncertain whether your colleague says that he will come on Tuesday or on Thursday (between 2.00 pm and 4.00 pm), but you have got a preference for him saying Tuesday.

2.3 The Update Mechanism

Definition 2.9 Given a pd-model $M = (W, \{\sim_j; j \in G\}, \{P_j; j \in G\}, V, w_0)$ and a generic action model $\mathbf{\Sigma} = (\Sigma, S, \{\sim_j; j \in G\}, P_j, \{P_j^\Gamma; \Gamma$ is a m.c.subset of S and $j \in G\}, \{Post_\sigma; \sigma \in \Sigma\}, \sigma_0)$, we define their update product to be the pd-model $M \otimes \mathbf{\Sigma} = (W \otimes \Sigma, \{\sim'_j; j \in G\}, \{P'_j; j \in G\}, V', w'_0)$, where:

1. $W \otimes \Sigma = \{(w, \sigma) \in W \times \Sigma; P_j^w(\sigma) > 0\}$.
2. $(w, \sigma) \sim'_j (v, \tau)$ iff $w \sim_j v$ and $\sigma \sim_j \tau$.
3. We set

   $$P'_j(\sigma) = \frac{P_j(\sigma).P_j^{[w]_j}(\sigma)}{\sum\{P_j(\tau).P_j^{[w]_j}(\tau); \sigma \sim_j \tau\}} \text{ where } P_j^{[w]_j}(\sigma) = \sum\{P_j(v).P_j^v(\sigma); w \sim_j v\}.$$

   Then

   $$P'_j(w, \sigma) = \frac{P_j(w)}{\sum\{P_j(v); w \sim_j v \text{ and } P_j^v(\sigma) > 0\}}.P'_j(\sigma).$$

4. $V'(p) = \{(w, \sigma) \in W \otimes \Sigma; M, w \models Post_\sigma(p)\}$.
5. $w'_0 = (w_0, \sigma_0)$.

$\triangleleft$

Intuitive Interpretation and Motivations.

Items 1 and 5 As in BMS, in the new model we consider all the possible worlds (w, σ) resulting from the performance of the possible action σ in the possible world w, granted that this action σ can physically take place in w (i.e. $P_j^w(\sigma) > 0$).

Item 2 As in BMS, the uncertainty relations $\sim_j$ for the pd-model and the generic action model are independent one from another. This independence allows us to 'multiply' these uncertainties to compute the new uncertainty relation.

Item 3. We want to determine $P'_j(w, \sigma) = P_j(W \cap A)$, where W stands for 'we were in world w before σ occurred' and A for 'action σ has occurred'. More formally, in the probability space $[(w, \sigma)]_j := \{(v, \tau); (v, \tau) \sim_j (w, \sigma)\}$, W stands for $\{(w, \tau); \tau \sim_j \sigma\}$ and A for $\{(v, \sigma); v \sim_j w\}$ and we can check that $W \cap A = \{(w, \sigma)\}$. But of course to determine these probabilities we have to rely only on M and $\mathbf{\Sigma}$.
Probability theory tells us that

$$P_j(W \cap A) = P_j(W|A).P_j(A).$$

We first determine $P_j(W|A)$, that is to say the probability that j was in world w *given* the extra assumption that action σ occurred in this world. We reasonably claim

$$P_j(W|A) = \frac{P_j(w)}{\sum\{P_j(v); w \sim_j v \text{ and } P_j^v(\sigma) > 0\}}.$$

That is to say, we *conditionalize* the probability of w for j (i.e. $P_j(w)$) to the worlds where the action σ took place and that may correspond for j to the actual world w (i.e. $\{v; w \sim_j v \text{ and } P_j^v(\sigma) > 0\}$). That is the way it would be done in classical probability theory. The intuition behind it is that we now possess the extra piece of information that σ occurred in w. So the worlds indistinguishable from w where the action σ did *not* occur do not play a role anymore for the determination of the probability of w. We can then get rid of them and conditionalize on the remaining relevant worlds.

It remains to determine $P_j(A)$, which we also denote $P'_j(\sigma)$; that is to say the probability for j that σ has occurred. We claim that

$$P'_j(\sigma) = P_1.P_2$$

where P_1 is the probability for j that σ has actually occurred, determined on the sole basis of her *perception* and *observation* of the action; and P_2 is the probability that j would have expected σ to happen determined on the sole basis of her epistemic state (i.e. her beliefs) in world w. Because P_1 and P_2 are independent, we simply multiply them to get $P'_j(\sigma)$.

By the very definition of $P_j(\sigma)$ (see Sect. 2.1), $P_1 = P_j(\sigma)$.

As for P_2, agent j's epistemic state in w is represented by the set of worlds $[w]_j := \{v; v \sim_j w\}$. So she could expect σ to happen in any of these worlds, each time with probability $P_j^v(\sigma)$. We might be tempted to take the average of them: $P_2 = \frac{\sum\{P_j^v(\sigma); v \sim_j w\}}{n}$, where n is the number of worlds in $[w]_j$. But we have more information than that on j's epistemic state. j does not know in which world of $[w]_j$ she is, but she has a preference among them, which is expressed by P_j. So we can refine our expression above and take the *center of mass* (or barycenter) of the $P_j^v(\sigma)$s balanced respectively by the weights $P_j(v)$s (whose sum equals 1), instead of taking roughly the average (which is actually also a center of mass but with weights $\frac{1}{n}$). We get $P_2 = P_j^{[w]_j}(\sigma) = \sum\{P_j(v).P_j^v(\sigma); v \sim_j w\}$. (Note that this expression could also be viewed as an application of a theorem of conditional probability if we rewrote $P_j^v(\sigma)$ $P_j(\sigma|v)$.)

Finally, we normalise $P_j'(\sigma)$ on the set of actions $[\sigma]_j$ to get a probabilistic space. We get

$$P_j(A) = P_j'(\sigma) = \frac{P_j(\sigma).P_j^{[w]_j}(\sigma)}{\sum\{P_j(\tau).P_j^{[w]_j}(\tau); \sigma \sim_j \tau\}} \text{ where } P_j^{[w]_j}(\sigma) = \sum\{P_j(v).P_j^v(\sigma); w \sim_j v\}.$$

Finally, we can easily check that the sum of $P_j'(w,\sigma)$ on $[(w,\sigma)]_j$ is equal to 1.

Item 4. Intuitively, this formula says that a fact p is true after the performance of σ in w iff the necessary and sufficient condition for p to be true after σ was satisfied in w before the action occurred.[4]

Example 2.10 [urn example] Assume now that I just drew a (white) ball from the urn and put it in my pocket, action depicted in Fig.3. Then because you did not have any particular preference about the distribution of the urn, you would expect that I draw a black ball or a white ball with equal probability. That is indeed the case:

$$P_Y^{[w_0]_Y}(\sigma) = \sum\{P_Y(w_i).P_Y^{w_i}(\sigma); w_i \sim_j w_0\} = \sum\{\tfrac{1}{n+1}.\tfrac{i}{n}; i = 0..n\} = \tfrac{1}{2} = P_Y^{[w_0]_Y}(\tau).$$

Independently from that, your perception of the action did not provide you any reason to prefer σ over τ ($P_Y(\sigma) = P_Y(\tau)$). So, in the end you should believe equally that I drew a black ball or a white ball, and this is indeed the case:

$$P_Y'(\sigma) = P_Y'(\tau).$$

If we perform the full update mechanism, then we get the pd-model of Fig.5. In this model all the worlds are equally probable for you and note that there cannot be n black balls in the urn (p_n) since one of them has been withdrawn.

[4] In a previous version of this paper ([Auc05b]), the solution of determining which propositional facts were true after an update was very similar to Reiter's solution of the frame problem ([Reit01]). It turns out that the solution we propose here is equivalent to the previous one and was first proposed by De Lavalette ([Ren 04]) and Kooi ([Kooi05]). Indeed, in [Auc05b] we used two operators $Post_\sigma^+$ and $Post_\sigma^-$ but in the update mechanism the necessary and sufficient condition for p to be true after σ was $Post_\sigma^+(p) \vee (p \wedge Post_\sigma^-(p))$ which is shortened here to $Post_\sigma(p)$.

$$
\begin{tikzcd}
\boxed{(w_1,\sigma):p_0} \arrow[r, dash, "B"] \arrow[d, dash, "B"] & \boxed{(w_2,\sigma):p_1} \arrow[r, dashed, dash, "B"] \arrow[d, dash, "B"] & \boxed{(w_k,\sigma):p_{k-1}} \arrow[r, dash, "B"] \arrow[d, dash, "B"] & \boxed{(w_{k+1},\sigma):p_{k+1}} \arrow[r, dashed, dash, "B"] \arrow[d, dash, "B"] & \boxed{(w_n,\sigma):p_{n-1}} \arrow[d, dash, "B"] \\
\boxed{\boxed{(w_0,\tau):p_0}} \arrow[r, dash, "B"] & \boxed{(w_1,\tau):p_1} \arrow[r, dashed, dash, "B"] & \boxed{(w_{k-1},\tau):p_{k-1}} \arrow[r, dash, "B"] & \boxed{(w_k,\tau):p_k} \arrow[r, dashed, dash, "B"] & \boxed{(w_{n-1},\tau):p_{n-1}}
\end{tikzcd}
$$

$P_Y(x) = \frac{1}{2n}$ for all worlds x

Figure 5: situation of the urn Example after I drew a white ball and put it in my pocket.

Now consider another scenario where this time you initially believe that there are more black balls than white balls (for example you believed me when I told you so at the beginning). This can be modelled initially by assigning the probabilities $P_Y(w_i) = \varepsilon$ for $i = 0, .., k$ and $P_Y(w_i) = \frac{1}{k}$ for $i = k+1, .., n$ to the worlds of the model depicted in Fig.1 (recall that there are $n = 2.k$ balls). Then, if I draw a ball from the urn and we compute again the probabilities of the actions σ and τ, we get what we expect:

$$P'_Y(\sigma) = P_Y^{[w_0]_Y}(\sigma) = \sum\{\varepsilon.\tfrac{i}{n}; i = 0..n\} + \sum\{\tfrac{1}{k}.\tfrac{i}{n}; i = k+1..n\} = \tfrac{3}{4} + \tfrac{1}{4.k} > \tfrac{1}{4} - \tfrac{1}{4.k} = P_Y^{[w_0]_Y}(\tau) = P'_Y(\tau).$$

Example 2.11 [answering machine example] Now that you have heard the message, you update your representation of the world with this new information. We are not going to perform the full update and display the new pd-model, but rather just concentrate on how you update your action probabilities $P'_Y(\sigma)$ and $P'_Y(\tau)$.

After this update, the probability $P'_Y(\sigma)$ for you that your colleague said that he will come on Tuesday is a combination of : (1) how much you would have expected him to say so, based on what you knew and believed of the situation, and (2) what you actually recognized and perceived from the answering machine.

The first value (1) is $P_Y^{[w]_Y}(\sigma) = \sum\{P_Y(v).P_Y^v(\sigma); v \sim_Y w\}$, and the second (2) is $P_Y(\sigma)$. We get $P'_Y(\sigma) = \frac{12}{29} < \frac{17}{29} = P'_Y(\tau)$. The important thing to note here is that σ has become less probable than τ for you: $P'_Y(\sigma) < P'_Y(\tau)$ while before the update $P_Y(\sigma) > P_Y(\tau)$. On the one hand, this is due to the fact that the probability that you heard your colleague saying that he would come on Tuesday has decreased due to your lower expectation of him to say so: $P_Y^{[w]_Y}(\sigma) = \frac{8}{25} < \frac{3}{5} = P_Y(\sigma)$; expectation which is based on your belief that he is busy on Tuesday because he has got a lecture ($P_Y(p) = \frac{4}{5}$). On the other hand, this is due to the fact that the probability that you heard your colleague saying that he would come on Thursday has increased due to your higher expectation of him to say so: $P_Y(\tau) = \frac{2}{5} < \frac{17}{25} = P_Y^{[w]_Y}(\tau)$).

Now consider a second scenario where this time you believe with a higher probability that he has got a lecture on Thursday than on Tuesday ($P_Y(w) = \frac{1}{5}$ and $P_Y(v) = \frac{4}{5}$ in

$$
\begin{tikzcd}
\boxed{\boxed{(w,\sigma):p}} \arrow[r, dash, "Y"] \arrow[d, dash, "Y"] & \boxed{(w,\tau):p} \arrow[d, dash, "Y"] \\
\boxed{(v,\sigma):\neg p} \arrow[r, dash, "Y"] & \boxed{(v,\tau):\neg p}
\end{tikzcd}
$$

$$P_Y(w,\sigma) = 4\varepsilon^2,\ P_Y(w,\tau) = \varepsilon,\ P_Y(v,\sigma) = 4\varepsilon,\ P_Y(v,\tau) = 1$$

Figure 6: Situation after you believed that your colleague's lecture is on Thursday and you heard him saying that he would come on Thursday.

$$\boxed{\boxed{\mu}}$$
$$P_Y(\mu) = 1$$
$$S = p, \neg p$$
$$P_Y^p(\mu) = 1, P_Y^{\neg p}(\mu) = 0$$
$$Post_\sigma(p) = p.$$

Figure 7: Your colleague announces you that his lecture is on Tuesday.

Fig.2). This time your belief that you heard him saying that he will come on Tuesday (resp. Thursday) is strengthened (resp. weakened) by your independent expectation of him to say so: $P'_Y(\sigma) > P_Y(\sigma) > P_Y(\tau) > P'_Y(\tau)$.

Example 2.12 [answering machine example 2] In this variant of the answering machine example, we are going to show the usefulness of infinitesimals and show an example of belief revision.

Basically, we consider the same situation, except that your beliefs are different. This time you are convinced that your colleague's lecture is on Thursday and you are also convinced that you heard him saying that he would come on Thursday. Formally, everything remains the same except that now in Fig.2 $P_Y(w) = \varepsilon$ and $P_Y(v) = 1$, and in Fig.4 $P_Y(\sigma) = \varepsilon$ and $P_Y(\tau) = 1$. If we apply the full update mechanism with these new parameters we get the model depicted in Fig.6. In this model you are convinced that he has got a lecture on Thursday and that he said he would come on Thursday (world (v,τ)). Moreover, you would be surprised (with intensity $5.\varepsilon$) if somebody claimed to you that your colleague has got his lecture on Tuesday *or* he said he will come on Tuesday (worlds (w,τ), (v,σ) and (w,σ) and we recall that $5.\varepsilon + 4.\varepsilon^2 = 5.\varepsilon$). But you would be much more surprised (with intensity ε^2) if somebody claimed to you that he has got his lecture on Tuesday *and* he said he will come on Tuesday (world (w,σ)), because that contradicts twice your original convictions.

Now if your colleague announces you that he has got his lecture on Tuesday, then you

$$\boxed{\boxed{((w,\sigma),\mu):p}} \overset{Y}{\text{———}} \boxed{((w,\tau),\mu):p}$$
$$P_Y((w,\sigma),\mu) = 4.\varepsilon, P_Y((w,\tau),\mu) = 1$$

Figure 8: situation after your colleague announced his lecture is on Tuesday and you then revised your beliefs.

will have to revise your beliefs. This action which is a public announcement is depicted in Fig.7. The resulting model is depicted in Fig.8. In this model, you now believe that he has got his lecture on Tuesday, but you are still convinced that your colleague said that he will come on Thursday because no new information has contradicted this. This is made possible thanks to the grading of surprising worlds by infinitesimal. What happened during this revision process is that the least surprising world where p is true became the conceived world.

3 Comparisons.

Structure of the static part and (its relevance for the) comparison with the AGM postulates.

There are several proposals in the literature which cumulate features of ranking theories (Spohn-type or possibility theories) and probability: for example generalized qualitative probability ([Lehm96]), lexicographic probability ([Blume Al 91]), big-stepped probabilities ([DubFarg04]) and cumulative algebra ([Wey94]). We showed in Theorem 1.6. that our structure $\mathbb{V}$ is actually an extension of a structure $\mathbb{V}'$ isomorphic to a cumulative algebra (notion introduced by Weydert). So Weydert's comparisons with Spohn's theory ([Spohn90])and possibility theories ([DubPrad91]) are still valid here. In particular the ranking $\mathbb{V}^0$ associated to $\mathbb{V}'$ defined in the appendix determines a *global* gradation of worlds similar in spirit to Spohn's degrees of disbelief or possibility degrees in possibility theory. (The global degree of a world w can be defined by $\alpha((P_j(w))^0)$ (see appendix for the notations used).) So, more precisely our conceived worlds correspond to Spohn's worlds of plausibility 0 and our surprising worlds correspond to Spohn's worlds of plausibility strictly greater than 0; moreover our global degrees corresponds to Spohn's plausibility degrees (although the order has to be reversed). Likewise with possibility theory. In the picture of Sect.2.1., the worlds with a same global degree are located between two adjacent concentric spheres. But among the worlds of a same global degree exists also a *local* gradation corresponding to the usual order relation. So for example, in the picture we have the local gradation $4.\varepsilon^2 > 3.\varepsilon^2$, even if the worlds with these probabilities have the same global degree.

As we saw in Sect.2.1. this formalism allows an accurate and rich account of an epistemic state. Now we are going to see that its duality (local and global aspects) enables also a fine-grained account of its dynamics as well.

Remark 3.1 In the literature (including myself in [Auc05a]), one often considers degrees of possibility/plausibility (present in possibility theory and Spohn's theory) and probabilities as two different means to represent and tackle the same kind of information. However, as it is stressed in this paper, for us they are meant to model two related but different kinds of information. In our sense, the first rather corresponds to degrees of potential surprise about facts absent from the agent's mind. The second rather corresponds to degrees of belief or acceptance about facts present (or accessible) in the agent's mind (which can be a knowledge base for example). The same distinction is also present in [Gard88].

Our structure $\mathbb{V}$ is richer than the cumulative algebra $\mathbb{V}'$ because it allows its elements to have multiplicative inverse. This feature turns out to be quite useful in a dynamic setting because it allows conditionalization and in particular belief revision.

Theorem 3.2 *If, as in the AGM theory, we restrict our attention to the single agent case and thus also to propositional beliefs, then in the case of a public announcement our update mechanism satisfies the 8 AGM postulates for belief revision.*

(The formal proof of this informal theorem is in the appendix and the proof needs to introduce some other natural assumptions.)

What actually occurs in the revision process by a formula φ, which is false in every conceived worlds, is that the surprising worlds where φ is true and which have the least *global* degree become the conceived worlds. More interestingly, the *local* structure of these surprising worlds remains the same, that is to say their relative order of probabilities is the same before and after the revision. So we see here that the richness of our formalism enables a fine-grained account of belief revision which is absent in the literature.

Comparison with Kooi's system.

Kooi's dynamic probabilistic system (see [Kooi03]), based on the static approach by Fagin and Halpern in [FH94], does not make any particular assumption about the relation between probability and epistemic accessibility relations (these different assumptions are explored in [FH94]). So, unlike us, probability is not meant only to model the agents' epistemic states. In that respect, his probability functions are defined relatively to each world. Moreover he only deals with public announcement which he does not assume to be truthful. For this reason and in this particular case our update mechanism is a bit different from his. The worlds of his initial model are the same as in his updated model, only the accessibility relations and probability distributions are changed, depending on whether or not the probability of the formula announced is null in the initial model. However, our probabilistic update rule in this particular case boils down to the same as his for the worlds where the probability of the formula announced is not null. Finally, he does not consider actions changing facts (he tackles this topic independently in [Kooi05]).

Comparison with van Benthem's system.

van Benthem's system (see [vBen03]) is similar to ours in its spirit and goals. However, his representation of an agent's epistemic state, based also on an equivalence relation $\sim_j$ and probability measures P_j, does not resort to infinitesimals and thus can neither express degrees of potential surprise nor allow belief revision. He does not deal with actions changing facts as well (although this is not the goal of his paper). More importantly, he does not show formally that there is an interplay when we interpret an action between what we actually perceive from the action and what we would have expected it to be. In that respect, he does not introduce the probabilities $P_j^{[w]_j}(\sigma)$ and $P_j(\sigma)$ but only a single $P_j^w(\sigma)$. Hence, his probabilistic update rule is completely different. The intended interpretation of his $P_j^w(\sigma)$ seems also to be different from ours if we refer to his example. Indeed, in our system we have the condition that for any possible action σ and world w, $\sum\{P_j^w(\tau); \tau \sim_j \sigma\} = 1$. So in our system we should have $P_j^w(\text{Q opens 3}) = 1$ for any world w because this public action can be distinguished by the agent from the other action "Q opens door 2" when it actually happens. However he sets $P_j^w(\text{Q opens 3}) = \frac{1}{2}$. This suggests a temporal difference in the determination of $P_j^w(\sigma)$ between the two definitions. Moreover he still uses preconditions which are generalized here to P_j^w (although he also suggests a different generalization of them). Anyway, his discussion and comparison with the Bayesian setting in his Sect.5 is still valid here.

Comparison with the situation calculus of Bacchus, Halpern and Levesque.

Their system (see [BaHaLe98]) can be viewed as the counterpart of van Benthem's system in the situation calculus except that they deal as well with actions changing facts. Their probabilistic update rule is also the same as van Benthem's (modulo normalisation). So what was said just above applies here two. In particular, the logical dynamics present in the interpretation of an action are not explored.

Comparison with the Observation Systems of Boutilier, Friedman and Halpern.

Their system (see [BoutAl98]) deals with noisy observations. Their approach is semantically driven like us. However they use a different framework called observation system based on the notion of (ranked) interpreted system ([FrieHal97]). On the one hand their system is more general because it incorporates the notion of time and a ranking of evolutions of states over time (called runs). On the other hand it is a single agent system and the only actions they consider are noisy observations (which do not change facts of the situation). An advantage of our system is its versatility because we can represent many kinds of actions. In that respect, their noisy observations can be modelled in our framework using two indistinguishable possible actions, the first corresponding to a truthful observation and the second to an erroneous one. Then by a suitable choice of probabilities we can express, as they do, that the observation is "credible" for example. However, their framework seems to enable them to characterize formally more types of noisy observations. Finally, because we do not introduce the notions of time and history, our framework is rather comparable to a particular case of their system called *Markovian* observation system. But nothing precludes us to introduce these notions as an extension of our system.

4 Conclusion

In order to represent with most accuracy the agents' epistemic state, we have introduced a rich formalism based on hyperreal numbers (and which is an extension of Weydert's cumulative algebra). Our epistemic state representation includes both degrees of beliefs expressed by a subjective probability and degrees of potential surprise expressed by infinitesimals. We have seen that the richness of this formalism enabled genuine belief revision thanks to the existence of infinitesimal (and multiplicative inverse). By a closer look at this revision process, we could even notice some interesting and meaningful patterns due to the dual aspect (local and global) of this formalism. So, our system indirectly offers a new (probabilistic) approach to belief revision.

But other important logical dynamics were studied, namely the ones present in the process of interpreting an action. Starting from the observation that this interpretation hinges on two features, the actual perception of the action and our expectation of it to happen, we have proposed a way to model this phenomenon. Note that in a sense our approach complements and reverses the classical view whereby only our interpretation of actions affects our beliefs and not the other way around, as in belief revision theory. For sake of generality, we have also taken into account in this system actions that change facts of a situation.

Our system is semantically driven and it would be interesting to look for a completeness result, and in particular for reduction axioms (for that we will surely need to define a richer language). However this system can be of use as it is in several areas. First, in game theory where the kinds of phenomenon we studied are quite current in games. Second, in psychology if we want to devise realistic formal models of belief change. And why not in artificial intelligence since the logical dynamics we modelled are the hallmark of rational and efficient reasoners.

References

[Auc05a] G. Aucher. A combined System for Update Logic and Belief Revision. In M. W. Barley and N. Kasabov (Eds.): *PRIMA 2004*, LNAI 3371 Pages 1-17, 2005.

[Auc05b] G. Aucher. How our beliefs contribute to interpret actions. In M. Pechoucek, P. Petta, L; Z; Varga (Eds.): *CEEMAS 2005*, LNAI 3690 Pages 276 - 286, 2005.

[Adams75] E. W. Adams. The Logic of Conditionals. In *Synthese Library* Volume 86, D. Reidel, Dordrecht, Netherlands, 1975.

[BaHaLe98] F. Bacchus, H. Levesque and Y. Halpern. Reasoning about noisy sensors in the situation calculus. In *Artificial Intelligence* 111, 1999, pp. 131-169.

[Blume Al 91] L. Blume, A. Brandenburger, E. Dekel. Lexicographic Probabilities and Choice Under Uncertainty, in *Econometrica*, Vol. 59, No. 1 (Jan., 1991) , pp. 61-79

[BMS04] A. Baltag, L.S. Moss,and S. Solecki. Logic for epistemic program. In *Synthese* Volume 139(2) Pages: 165 - 224.

[BoutAl98] C. Boutilier, N. Friedman and J. Halpern. Belief Revision with Unreliable Observations. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence* (AAAI-98), pp 127-134, 1998.

[DubFarg04] D. Dubois, H. Fargier. A Unified Framework for Order-of-Magnitude confidence Relations. *Twentieth Conference in Artificial Conference.* M. Chickering, J. Halpern Eds., pp 138-145, Banff, Canada, July 2004.

[DubPrad91] D. Dubois, H. Prade. Possibilistic logic, preferential models and related issues. In *Proceedings of the 12th International Conference on Artificial Intelligence*, Morgan Kaufmann, 1991.

[FH94] R. Fagin and J. Y. Halpern. Reasoning about knowledge and probability. In *Journal of the ACM* (JACM) Volume 41(2) Pages: 340 - 367.

[FrieHal97] N. Friedman and J. Halpern. Modeling Beliefs in Dynamic Systems, part I:Foundations. *Artificial Intelligence*, 95:257-316, 1997.

[Gard88] P. Gardenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, Bradford Books, MIT Press, Cambridge, Mass., 1988.

[Keis86] H. J. Keisler. *Elementary Calculus: An Approach Using Infinitesimal.* Prindle, Weber and Schmidt (eds), 1986. Online edition on the website http://www.math.wisc.edu/ keisler/calc.html .

[Kooi03] B.P. Kooi (2003). Probabilistic Dynamic Epistemic Logic. In *Journal of Logic, Language and Information* Volume 12(4) Pages: 381-408.

[Kooi05] B. Kooi. *As the World Turns: On The Logic of Public Update.* Manuscript. 2005.

[Lehm96] D. Lehmann. Generalized Qualitative Probability: Savage revisited, *Twelfth Conference on Uncertainty in Artificial Intelligence*, E. Horvitz and F. Jensen eds., Morgan Kaufmann, pp. 381-388 Portland, Oregon, August 1996

[Len78] W. Lenzen. *Recent Work in Epistemic Logic.* Amsterdam (North Holland Publ. Comp.), 1978 (Acta Philosophica Fennica 30).

[Reit01] R. Reiter. *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems*, MIT Press, 2001.

[Ren 04] Renardel De Lavalette G. R., Changing modalities, Journal of Logic and Computation, vol. 14, num. 2, 2004, p. 251275.

[Spohn90] W. Spohn. A general non-probability theory of inductive reasoning. In R. D. Schachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer, editors, *Uncertainty in artificial intelligence 4*, pages149-15. Norht-Holland, Amsterdam, 1990.

[vBen03] J. van Benthem. Conditional probability meets update logic. In *Journal of Logic, Language and Information* Volume 12(4) Pages: 409 - 421.

[Wey94] E. Weydert. General belief measures. In *tenth Conference on Uncertainty in Artificial Intelligence.* Morgan Kaufmann, 1994.

A Proof of Theorem 1.2., Lemma 1.4. and Theorem 1.5.

Theorem A.1 (Theorem 1.2.) *The quotient structure* $\mathbb{V} = ({}^*\mathbb{R}^+/_{\approx}, \overline{+}, \overline{\cdot})$ *is a semi-field.*

Proof. We only need to prove that $\overline{+}$ and $\overline{\cdot}$ are well defined, since the rest is standard and straightforward.

Assume $\overline{x} = \overline{x'}$ and $\overline{y} = \overline{y'}$. We have to show $\overline{x}\overline{+}\overline{y} = \overline{x'}\overline{+}\overline{y'}$ and $\overline{x}\overline{\cdot}\overline{y} = \overline{x'}\overline{\cdot}\overline{y'}$.

- First, let us show that $\overline{x}\overline{+}\overline{y} = \overline{x'}\overline{+}\overline{y'}$.

  Assume $\overline{x} = 0$ (similar proof for $\overline{y} = 0$). Then $x = x' = 0$. In that case $\overline{x}\overline{+}\overline{y} = \overline{x+y} = \overline{y} = \overline{y'} = \overline{0+y'} = \overline{x'+y'} = \overline{x'}\overline{+}\overline{y'}$.
  Assume $\overline{x} \neq 0$ and $\overline{y} \neq 0$. Then x, x', y, y' are all different from 0. So $x+y \neq 0$ because $x, y \geq 0$.

  $\overline{x}\overline{+}\overline{y} = \overline{x'}\overline{+}\overline{y'}$ iff
  $\overline{x+y} = \overline{x'+y'}$ iff
  $St(\frac{x'+y'}{x+y}) = 1$ because $x+y \neq 0$ (see above) iff
  $St(\frac{x}{x+y}) + St(\frac{y'}{x+y}) = 1$ iff
  $St(\frac{x'}{x} \cdot \frac{1}{1+\frac{y}{x}}) + St(\frac{y'}{y} \cdot \frac{1}{1+\frac{x}{y}}) = 1$ iff
  $St(\frac{x'}{x}).St(\frac{1}{1+\frac{y}{x}}) + St(\frac{y'}{y}).St(\frac{1}{1+\frac{x}{y}}) = 1$ iff
  $St(\frac{1}{1+\frac{y}{x}}) + St(\frac{1}{1+\frac{x}{y}}) = 1$ because $St(\frac{x'}{x}) = St(\frac{y'}{y}) = 1$ iff
  $\frac{1}{1+St(\frac{y}{x})} + \frac{1}{1+St(\frac{x}{y})} = 1$ iff
  $\frac{1}{1+St(\frac{y}{x})} + \frac{1}{1+\frac{1}{St(\frac{y}{x})}} = 1$ which is true.

- Now, let us show that $\overline{x}\overline{.}\overline{y} = \overline{x'}\overline{.}\overline{y'}$.

  Assume $\overline{x} = 0$ (similar proof for $\overline{y} = 0$). Then $x = x' = 0$ and the equality is fulfilled.

  Assume $\overline{x} \neq 0$ and $\overline{y} \neq 0$. Then x, x', y, y' are different from 0. So $x.y \neq 0$.

  $\overline{x}\overline{.}\overline{y} = \overline{x'}\overline{.}\overline{y'}$ iff
  $\overline{x.y} = \overline{x'.y'}$ iff
  $St(\frac{x'.y'}{x.y}) = 1$ because $x.y \neq 0$ iff
  $St(\frac{x'}{x}).St(\frac{y'}{y}) = 1$ iff
  $1.1 = 1$ which is true.

QED

Lemma A.2 (Lemma 1.4.) *If $\overline{x} \lessapprox \overline{y}$ then for all $x' \in \overline{x}$ and all $y' \in \overline{y}$, $x' \leq y'$*

Proof.

1. Assume $\overline{x} \neq 0$ and $\overline{y} \neq 0$ (then x, y are different from 0).
   (a) If $\overline{x} = \overline{y}$ then we have the result.
   (b) If $\overline{x} \neq \overline{y}$ then by definition there are $x_0 \in \overline{x}$ and $y_0 \in \overline{y}$ such that $x_0 < y_0$.
   Assume there are $x' \in \overline{x}, y' \in \overline{y}$ such that $x' > y'$

- Assume that either $x_0 \leq y' \leq x' \leq y_0$ or $x_0 \leq y' \leq y_0 \leq x'$ or $x_0 \leq y_0 \leq y' \leq x'$.

  Then $x_0 \leq y' \leq x'$, so
  $\frac{x_0}{x_0} \leq \frac{y'}{x_0} \leq \frac{x'}{x_0}$ because $x_0 \neq 0$, then
  $1 \leq St(\frac{y'}{x_0}) \leq St(\frac{x'}{x_0}) = 1$ because $x', x_0 \in \overline{x}$, then
  $St(\frac{y'}{x_0}) = 1$,
  $y' \in \overline{x_0} = \overline{x}$,
  $\overline{x} = \overline{y'} = \overline{y}$ which is impossible by assumption.

- Assume that either $y' \leq x_0 \leq x' \leq y_0$ or $y' \leq x' \leq x_0 \leq y_0$ or $y' \leq x_0 \leq y_0 \leq x'$.

  Then $y' \leq x_0 \leq y_0$, then
  $\frac{y'}{y_0} \leq \frac{x_0}{y_0} \leq 1$ because $y_0 \neq 0$, then
  $1 = St(\frac{y'}{y_0}) \leq St(\frac{x_0}{y_0}) \leq 1$, then
  $St(\frac{x_0}{y_0}) = 1$, then
  $\overline{x} = \overline{x_0} = \overline{y_0} = \overline{y}$ which is impossible by assumption.

So in all possible cases we reach a contradiction. This means that for all $x' \in \overline{x}$, all $y' \in \overline{y}$ $x' \leq y'$

2. (a) If $\overline{x} = 0$ and $\overline{y} \neq 0$, then
   $0 \lessapprox \overline{y}$ then there is $y_0 \in \overline{y}$ such that $0 < y_0$.

   Assume there is $y' \in \overline{y}$ such that $0 > y'$. Then
   $y' < 0 \leq y_0$
   $\frac{y'}{y_0} < 0 \leq \frac{y_0}{y_0} = 1$
   $1 = St(\frac{y'}{y_0}) \leq 0 \leq 1$ because $y', y_0 \in \overline{y}$
   i.e. $0 = 1$, which is counterintuitive.

   So for all $y' \in \overline{y}$, $0 \leq y'$.

   (b) if $\overline{y} = 0$ and $\overline{x} \neq \overline{0}$ then
   $\overline{x} \leq 0$, so there is $x_0 \in \overline{x}$ such that $x_0 < 0$.

   Assume there is $x' \in \overline{x}$ such that $x' > 0$.
   $x_0 \leq 0 < x'$
   $1 = \frac{x_0}{x_0} \leq 0 < \frac{x'}{x_0}$
   $1 \leq 0 \leq St(\frac{x'}{x_0}) = 1$ because $x', x_0 \in \overline{x}$
   i.e. $0 = 1$ which is again counterintuitive.

   So for all $x' \in \overline{x}$, $x' \leq 0$.

QED

Theorem A.3 (Theorem 1.5.) *The structure* $(\mathbb{V}, \lessapprox)$ *is an ordered semi-field.*

Proof. First we prove a lemma:

Lemma A.4 *$\lessapprox$ is a total order on $^*\mathbb{R}^+$ such that*

1. *if $\overline{x} \lessapprox \overline{y}$ then $\overline{x\overline{+}z} \lessapprox \overline{y\overline{+}z}$,*
2. *if $0 \lessapprox \overline{x}$ and $0 \lessapprox \overline{y}$ then $0 \lessapprox \overline{x\bar{.}y}$.*

Proof. Follows easily from lemma 1.4 and the fact that $\leq$ is a total order on $^*\mathbb{R}^+$ satisfying also conditions 1 and 2 above. QED

The proof then follows easily from the lemma above and theorem 1.2. QED

B Proof of Theorem 1.6.

First we are going to define the ranking $\mathbb{V}^0$ associated to $\mathbb{V}'$. To do so we first introduce an equivalence relation on $\mathbb{V}'$.

Definition B.1 Let $x, y \in \mathbb{V}'$, we set

$$x \approx^0 y \text{ iff } \begin{cases} \frac{x}{y} \text{ is real different from } 0 & \text{if } y \neq 0 \\ x = 0 & \text{if } y = 0. \end{cases}$$

We can easily check that $\approx^0$ is an equivalence relation on $\mathbb{V}'$. $\triangleleft$

Definition B.2 We define the ranking $\mathbb{V}^0 = (V^0, +^0, .^0, 0^0, 1^0, \leq^0)$ by the quotient structure of $\mathbb{V}'$ by the equivalence relation $\approx^0$. $\triangleleft$

We can easily check that $\mathbb{V}^0$ is well defined. Now we can prove theorem 1.6.

Theorem B.3 (Theorem 1.6.) *$\mathbb{V}'$ is isomorphic to the cumulative algebra with global structure $\mathbb{V}^0$ and local structure $\mathbb{R}^+$.*

Proof. We assume the validity of the axiom of choice and so the existence of a function $\alpha : \mathbb{V}^0 \to \mathbb{V}'$ which assigns to each element x^0 of $\mathbb{V}^0$ (which is a subset of $\mathbb{V}'$) an element of $\mathbb{V}'$ such that $\alpha(x^0) \in x^0$.

Now we can define an isomorphism $f : \mathbb{V}' \to \mathbb{H}(\mathbb{V}^0, \mathbb{R}^+)$ between $\mathbb{V}'$ and the cumulative algebra with global structure $\mathbb{V}^0$ and local structure $\mathbb{R}^+$ as follows: $f(x) = (x^0, \frac{x}{\alpha(x^0)})$. Its inverse isomorphism $g : \mathbb{H}(\mathbb{V}^0, \mathbb{R}^+) \to \mathbb{V}'$ is defined by $g(x^0, y) = \alpha(x^0).y$. QED

C Proof that our System Satisfies the 8 AGM Postulates

To check whether the AGM postulates are fulfilled, we first need to define, relatively to the single equivalence class $[w]_j$ of the single agent j, the belief set, the expanded belief set and the revised belief set. We will deal with propositional language as in the AGM theory. The type of generic action model we naturally consider for the update is a public announcement of a propositional formula φ, depicted in Fig.4.

Since we did not introduce the full dynamic language but only the static one, we just use a formal shortcut to deal in particular with public announcement.

$$\boxed{\boxed{\varphi!}}$$
$$S = \{\varphi, \neg\varphi\} \text{ and } P_j^{\{\varphi\}}(\sigma) = 1, P_j^{\{\neg\varphi\}}(\sigma) = 0 \text{ for the single agent } j.$$

Figure 9: generic action model for the action 'public announcement of the propositional formula φ'

Definition C.1 Let $M = (W, \sim_j, P_j, V, w_0)$ be a pd-model. Let φ, ψ be propositional formulas. We define

$M, w \models [\varphi!]\psi$ iff, if $M, w \models \varphi$ then $M(\varphi!), (w, \varphi!) \models \psi$,

where $M(\varphi!)$ is the pd-model defined by $M(\varphi!) := (W' := \{(w, \varphi!); M, w \models \varphi\}, \sim'_j := \{((w, \varphi!), (v, \varphi!)); w \sim_j v\}, \{P_j(w, \varphi!) = \frac{P_j(w)}{\sum\{P_j(v); v \sim_j w \text{ and } v \models \varphi\}}; (w, \varphi!) \in W'\}, V' := V, w'_0 := (w_0, \varphi!))$ $\lhd$

Definition C.2 Let $\mathcal{L}$ be the propositional language. For each equivalence class $[w]_j$ we associate a world v such that $w \sim_j v$ and $M, v \models \varphi$, and we define

- the belief set $K^{[w]_j} = \{\psi \in \mathcal{L}; M, v \models C_j\psi\}$,
- the revision of the belief set $K^{[w]_j}$ by φ, $K^{[w]_j} * \varphi = \{\psi \in \mathcal{L}; M, v \models [\varphi!]C_j\psi\}$,
- the expansion of the belief set $K^{[w]_j}$ by φ, $K^{[w]_j} + \varphi = \{\psi \in \mathcal{L}; M, v \models C_j(\varphi \to \psi)\}(= \{\psi \in \mathcal{L}; M, w \models C_j([\varphi!]\varphi \to \psi)\}$ because we use the propositional language).

$\lhd$

Theorem C.3 *(If $M, w \models P_j(\varphi) > 0$ then) * defined by $K^w * \varphi$ satisfies the 8 AGM postulates.*

In the theorem, the assumption within brackets is a natural one and $P_j(\varphi)$ can even be infinitesimal.

Proof.

(K*1) Yes, clearly.

(K*2) Yes, because we deal with propositional formulas which are persistent formulas (that is formulas which remain true after a public announcement if true beforehand).

(K*3) $M, v \models [\varphi!]C_j\psi$

$\Leftrightarrow \sum\{P_j(v', \varphi!); (v', \varphi!) \sim_j (v, \varphi!) \text{ and } M(\varphi!), (v', \varphi!) \models \psi\} = 1$

$\Leftrightarrow \sum\{\frac{P_j(v')}{\sum\{P_j(u); u \sim_j v' \text{ and } M, u \models \varphi\}}; v' \sim_j v \text{ and } M, v' \models \varphi \text{ and } M, v' \models \psi\} = 1$ because propositional formulas are persistent.

$\Leftrightarrow \sum\{P_j(v'); v' \sim_j v \text{ and } M, v' \models \varphi \text{ and } M, v' \models \psi\} = \sum\{P_j(v'); v' \sim_j v \text{ and } M, v' \models \varphi\}$

$\Rightarrow^* \sum\{P_j(v'); v' \sim_j v \text{ and } M, v' \models \varphi \to \psi\} = 1$ (1)

$\Leftrightarrow M, v \models C_j(\varphi \to \psi)$

(K*4) $\neg\varphi \notin K^{[w]_j}$

$\Leftrightarrow M, v \models \neg C_j \neg\varphi$

$\Leftrightarrow$ there is a conceived world v' such that $v \sim_j v'$ and $M, v' \models \varphi$ (H).

We have to prove the other direction of $\Rightarrow^*$.

Formula (1) tells us that for all conceived worlds v' such that $v' \sim_j v$, $M, v' \models \varphi \to \psi$. So,

$\sum\{P_j(v'); v' \sim_j v \text{ and } M, v' \models \varphi\}$

$= \sum\{P_j(v'); v' \sim_j v \text{ and } M, v' \models \varphi \text{ and } v' \text{ conceived }\}$ by (H), because if u is conceived and u' surprising then $P_j(u) + P_j(u') = P_j(u)$ (see Sect.1)

$= \sum\{P_j(v'); v' \sim_j v \text{ and } M, v' \models \varphi \text{ and } M, v' \models \psi \text{ and } v' \text{ conceived }\}$ by (1)

$= \sum\{P_j(v'); v' \sim_j v \text{ and } M, v' \models \varphi \text{ and } M, v' \models \psi\}$ by (H).

(K*5) It is fulfilled because $K^{[w]_j} * \varphi \neq K_\perp$.

Indeed, otherwise $M, v \models [\varphi!] C_j \perp$

but $\models \neg C_j \perp$, so $\models [\varphi!] \neg C_j \perp$.

Then $M, v \models [\varphi!] \perp$.

That is $M, v \models \neg\varphi$

which is wrong by assumption.

(K*6) Yes, clearly.

(K*7) We naturally assume moreover that $M, v \models \varphi'$. First note that $M, v \models [\varphi!] C_j(\varphi' \to \psi) \Leftrightarrow \varphi \in K^{[w]_j} * \varphi + \varphi'$

$M, v \models [\varphi \wedge \varphi'] C_j \psi$

$\Leftrightarrow$ if $M, v \models \varphi \wedge \varphi'$ then $M(\varphi \wedge \varphi'!), (v, \varphi \wedge \varphi'!) \models C_j \psi$

$\Leftrightarrow M(\varphi \wedge \varphi'!), (v, \varphi \wedge \varphi'!) \models C_j \varphi$ by assumption

$\Leftrightarrow \sum\{P_j(v', \varphi \wedge \varphi'!); v' \sim_j v \text{ and } M, v' \models \varphi \wedge \varphi' \text{ and } M, v' \models \psi\} = 1$

$\Leftrightarrow \sum\{\frac{P_j(v')}{\sum\{P_j(u); u \sim_j v' \text{ and } M, u \models \psi \wedge \varphi'\}}; v' \sim_j v \text{ and } M, v' \models \varphi \wedge \varphi' \text{ and } M, v' \models \varphi\} = 1$

$\Leftrightarrow \sum\{P_j(v'); v' \sim_j v \text{ and } M, v' \models \varphi \wedge \varphi' \text{ and } M, v' \models \varphi\} = \sum\{P_j(v'); v' \sim_j v \text{ and } M, v' \models \varphi \wedge \varphi'\}$

$\Leftrightarrow \sum\{\frac{P_j(v')}{\sum\{P_j(u); u \sim_j v' \text{ and } M, u \models \varphi\}}; v' \sim_j v \text{ and } M, v' \models \varphi \wedge \varphi' \text{ and } M, v' \models \psi\} = \sum\{\frac{P_j(v')}{\sum\{P_j(u); u \sim_j v' \text{ and } M, u \models \varphi\}}; v' \sim_j v \text{ and } M, v' \models \varphi \wedge \varphi'\}$

$\Leftrightarrow \sum\{P_j(v', \varphi!); (v', \varphi!) \sim_j (v, \varphi!) \text{ and } M, v' \models \varphi \wedge \varphi' \text{ and } M, v' \models \psi\} = \sum\{P_j(v', \varphi!); (v', \varphi!) \sim_j (v, \varphi!) \text{ and } M, v' \models \varphi \wedge \varphi'\}$

$\Leftrightarrow \sum\{P_j(v', \varphi!); (v', \varphi!) \sim_j (v, \varphi!) \text{ and } M(\varphi!), (v', \varphi!) \models \varphi' \text{ and } M(\varphi!), (v', \varphi!) \models \psi\} = \sum\{P_j(v', \varphi!); (v', \varphi!) \sim_j (v, \varphi!) \text{ and } M(\varphi!), (v', \varphi!) \models \varphi'\}$

$\Rightarrow^{*'} \sum\{P_j(v', \varphi!); (v', \varphi!) \sim_j (v, \varphi!) \text{ and } M(\varphi!), (v', \varphi!) \models \varphi' \to \psi\} = 1$ (1')

$\Leftrightarrow M, v \models [\varphi!] C_j(\varphi' \to \psi)$

$\Leftrightarrow \psi \in K^{[w]_j} * \varphi + \varphi'$

(K*8) $\neg\varphi' \notin K^{[w]_j} * \varphi$

$\Leftrightarrow M, v \models \neg[\varphi!]C_j\neg\varphi'$

$\Leftrightarrow M(\varphi!), (v, \varphi!) \models \neg C_j\neg\varphi'$

$\Leftrightarrow$ there is a conceived world $(v', \varphi!)$ such that $(v', \varphi!) \sim_j (v, \varphi!)$ and $M(\varphi!), (v', \varphi!) \models \varphi'$ (H')

We have to prove the other direction of $\Rightarrow^{*'}$. (1') tells us that for all conceived worlds $(v', \varphi!)$ such that $(v', \varphi!) \sim_j (v, \varphi!)$, $M(\varphi!), (v', \varphi!) \models \varphi' \to \psi$. So,

$\sum\{P_j(v', \varphi!); (v', \varphi!) \sim_j (v, \varphi!)$ and $M(\varphi!), (v', \varphi!) \models \psi\}$

$= \sum\{P_j(v', \varphi!); (v', \varphi!) \sim_j (v, \varphi!)$ and $M(\varphi!), (v', \varphi!) \models \psi$ and $(v', \varphi!)$ is conceived $\}$ by (H'), because if u is conceived and u' surprising then $P_j(u) + P_j(u') = P_j(u)$ (see Sect.1).

$= \sum\{P_j(v', \varphi!); (v', \varphi!) \sim_j (v, \varphi!)$ and $M(\varphi!), (v', \varphi!) \models \psi$ and $M(\varphi!), (v', \varphi!) \models \psi$ and $(v', \varphi!)$ conceived $\}$ by (1')

$= \sum\{P_j(v', \varphi!); (v', \varphi!) \sim_j (v, \varphi!)$ and $M(\varphi!), (v', \varphi!) \models \varphi'$ and $M(\varphi!), (v', \varphi!) \models \psi\}$ by (H').

So the other direction of $\Rightarrow^{*'}$ is proved.

QED