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

Guillaume Aucher

## HAL Id: inria-00556151
## https://hal.inria.fr/inria-00556151

Submitted on 1 Sep 2013

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

**Guillaume Aucher**

*Université Paul Sabatier - University of Otago*
*Postal Address: IRIT - Equipe LILaC*
*118 route de Narbonne*
*F-31062 Toulouse Cedex 9 (France)*

*aucher@irit.fr*

ABSTRACT. *In update logic as studied by Baltag, Moss, Solecki and van Benthem, little attention is paid to the interpretation of an action by an agent, 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 accurately 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.*

KEYWORDS: *probability, dynamic-epistemic logic, belief revision, update.*

The interpretation of an action by an agent 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 $n$ balls which are either black or white. If you *believe* that it is equally probable that there are $0,1,...,$ or $n$ black balls in the urn then you 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 expect with a higher probability that he draws a black ball rather than a white ball. We see in this example that your beliefs about the situation contribute actively to interpret the action: they determine the probability with which you expect the white-ball drawing to happen. But this expectation, determined by your beliefs about the situation, can often be balanced consciously or unconsciously by what you actually obtain by the pure perception of the action happening. 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 Thurs-

day. 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 consider more probable of having heard Tuesday than 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 (observation) of the action happening. Up to now, this kind of phenomenon, although very common in everyday life is not dealt with neither in update logic [BAL 04], [BEN 03], [AUC 05] nor in the situation calculus [BAC 99].

In order to represent the agent's 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 expressiveness 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 agent (and how much) by contradicting his/her beliefs.

In this paper we follow the approach of update logic as viewed by Baltag, Moss and Solecki (to which we will refer by the term BMS) [BAL 98]. For sake of generality, the actions we consider in this paper can also change the (propositional) facts of the situation[1]. As in BMS, we divide our task in three parts. Firstly, we propose a formalism called proba-doxastic (pd) model to represent how the actual world is perceived by the agent 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 agent (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 is the agent's representation of the world after the action represented by the above generic action model took place in the world represented by the above pd-model (Sect. 2.3). Our account will be presented for a single agent to highlight the main new ideas, but in Sect. 3 we will sketch some possible extensions to the multi-agent case. Finally we give some comparisons with the AGM postulates and relevant literature (Sect. 4).

NOTE 1. — All the proofs of lemmas and theorems of this paper are in the appendix. At the address *ftp://ftp.irit.fr/IRIT/LILAC/Aucher_jancl.pdf* can also be found a multi-agent version of this paper.                                                                         □

---

1. In the literature there is sometimes a distinction made between the terms "event" and "action", an action being a particular kind of event. In this paper the term action refers to the same general term as event.

## 1. Mathematical preliminaries

In our system, the probabilities of worlds and formulas will take values in a particular mathematical structure $(\mathbb{V}, \lesssim)$ (abusively noted $(\mathbb{V}, \leq)$) different from the real numbers, based on hyperreal numbers $(^*\mathbb{R}, \leq)$. (The approach in [ADA 75] uses them as well to give a probabilistic semantics to conditional logic.) In this section, we will briefly recall the main features of hyperreal numbers that will be useful in the sequel (for details see [KEI 86]). 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 $\varepsilon$, finite numbers are denoted $x$ and infinite numbers are denoted $\infty$. Note that an infinitesimal number is a finite number as well, that $\frac{1}{\varepsilon}$ is an infinite number and that $\frac{1}{\infty}$ is an infinitesimal number. Two hyperreal numbers $x$ and $y$ are said to be infinitely close to each other if their difference $x - y$ 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 $St(1 + \varepsilon) = 1$, $St(\varepsilon) = 0$.

Hyperreal numbers will be used to assign probabilities to facts (formulas). A fact is considered consciously probable by the agent when its probability is real. A fact would surprise the agent if (s)he learnt that it was true when its probability is infinitesimal. We want to refine the ordering given by the infinitesimals and introduce a global ranking among these potentially surprising facts. Indeed, a fact of probability $\varepsilon^2$ is infinitely more surprising than a fact of probability $\varepsilon$ and hence, for the agent, the importance of the former should be negligible compared to the importance of the latter. This can be done algebraically by approximating our expressions. More precisely, in case a hyperreal number $x$ is infinitely smaller than $y$, i.e. there is an infinitesimal $\varepsilon$ such that $x = \varepsilon.y$, then we want $y + x = y$. For example we want $1 + \varepsilon = 1$ (here $x = \varepsilon$ and $y = 1$), $\varepsilon + \varepsilon^2 = \varepsilon$ (here $x = \varepsilon^2$ and $y = \varepsilon$),... In other words, in case $x$ is *negligible* compared to $y$, then $y + x = y$. 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. 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 2. — *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 $^*\mathbb{R}^+$.*

For instance, we have $1 + \varepsilon \approx 1$ and $\varepsilon + \varepsilon^2 \approx \varepsilon$.

THEOREM 3. — *The quotient structure* $\mathbb{V} = ({}^*\mathbb{R}^+/_{\approx}, \overline{+}, \overline{.})$ *is a semi-field. (Elements of* $\mathbb{V}$, *being equivalence classes of* ${}^*\mathbb{R}^+$, *are classically denoted* $\overline{x}$. *And* $\overline{+}, \overline{.}$ *denote the quotient relations of* $+$ *and ..)*

Now we need to define the ordering relation $\lesssim$ on $\mathbb{V}$.

DEFINITION 4. — *We define a relation* $\lesssim$ *on* $\mathbb{V}$ *by*

$\overline{x} \lesssim \overline{y}$ *iff there are* $x \in \overline{x}$, $y \in \overline{y}$ *such that* $x \leq y$

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

LEMMA 5. — *If* $\overline{x} \lesssim \overline{y}$ *then for all* $x' \in \overline{x}$ *and all* $y' \in \overline{y}$, $x' \leq y'$

THEOREM 6. — *The structure* $(\mathbb{V}, \lesssim)$ *is an ordered semi-field.*

In the sequel, we will denote abusively $\overline{x}, \overline{+}, \overline{.}, \lesssim$ 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 $\{x \in \mathbb{V}; 0 \lesssim x \lesssim 1 \text{ and not } x \approx 0\}$.

We can then easily check that now we do have at our disposal the following identities: $1 + \varepsilon = 1, 0.6 + \varepsilon = 0.6, \varepsilon + \varepsilon^2 = \varepsilon, \ldots$

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 [WEY 94].

THEOREM 7. — *The set* $\mathbb{V}' = \{x \in \mathbb{V}; x \text{ is real or } x \text{ is infinitesimal }\}$ *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 8. — *A proba-doxastic model (pd-model) is a tuple* $M = (W, P, V, w_0)$ *where:*

*1)* $W$ *is a finite set of possible worlds;*

*2)* $w_0$ *is the possible world corresponding to the actual world;*

*3)* $P$ *is a probabilistic measure defined on* $\mathcal{P}(W)$ *which assigns to each world* $w \in W$ *a number in* $]0; 1]$;

*4)* $V$ *is a mapping called valuation which assigns to each propositional letter* $p$ *a subset* $V(p)$ *of* $W$.

***Intuitive interpretations.***

The possible worlds $W$ are determined by the modeler. One of them, $w_0$, corresponds to the actual world. Among these worlds $W$ there are some worlds that the agent conceives as potential candidates for the world in which she dwells, and some that she would be surprised to learn that they actually correspond to the world in which she dwells (whether this is true or false). The first ones are called *conceived worlds* and the second *surprising worlds*. The conceived worlds are assigned by $P$ a real value and the surprising worlds are assigned an infinitesimal value, both different from 0. For example, some people would be surprised if they learnt that some swans are black, although it is true. To model this situation, we introduce two worlds: one where all swans are white (world $w$) and one where some swans are not white (world $v$). So for these people the actual world $v$ is a surprising world, whereas the world $w$ is a conceived world.

Of course for the agent some (conceived) worlds are better candidates than others, and this is expressed by the probability value of the world: the larger the real probability value of the (conceived) world is, the more likely it is for the agent. But that is the same for the surprising worlds: the agent might be more surprised to learn about 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 was a surprising world for you). But you will be even more surprised if he tells you that he has cheated five times before. 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 the agent would be surprised by this world. Anyway, that is why we need to introduce hyperreal numbers: to express these degrees of potential surprise that cannot be expressed by a single number like 0 (which then becomes useless for us).

The agent does not think consciously that the surprising worlds 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 modelling from an objective viewpoint (see also Remark 17 in Sect. 2.3). Indeed they provide some information about the epistemic state of the agent: 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 if you learn that it is true. These worlds will moreover turn out to be very useful technically in case the agent has to revise her beliefs (see Sect. 3).

As $P$ is a probability measure defined on the $\sigma$-algebra $\mathcal{P}(W)$, by definition we have $\sum\{P(v); v \in W\} = 1$. Moreover, since $P$ takes its values in $\mathbb{V}'$ which is isomorphic to a cumulative algebra (Theorem 7), $P$ is in fact a cumulative measure in the sense of Weydert [WEY 94]. In Fig. 1 is depicted an example of a pd-model

(without any valuation). 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. The other circles correspond to the global ranking of surprising worlds that was motivated in Sect. 1. This global ranking is closely related to Weydert's notion of ranking and will be studied more in detail in Sect. 4. Note that both the sum of the probabilities of the conceived worlds and the global sum of all these worlds (conceived and surprising) are equal to 1 (see Sect. 1).
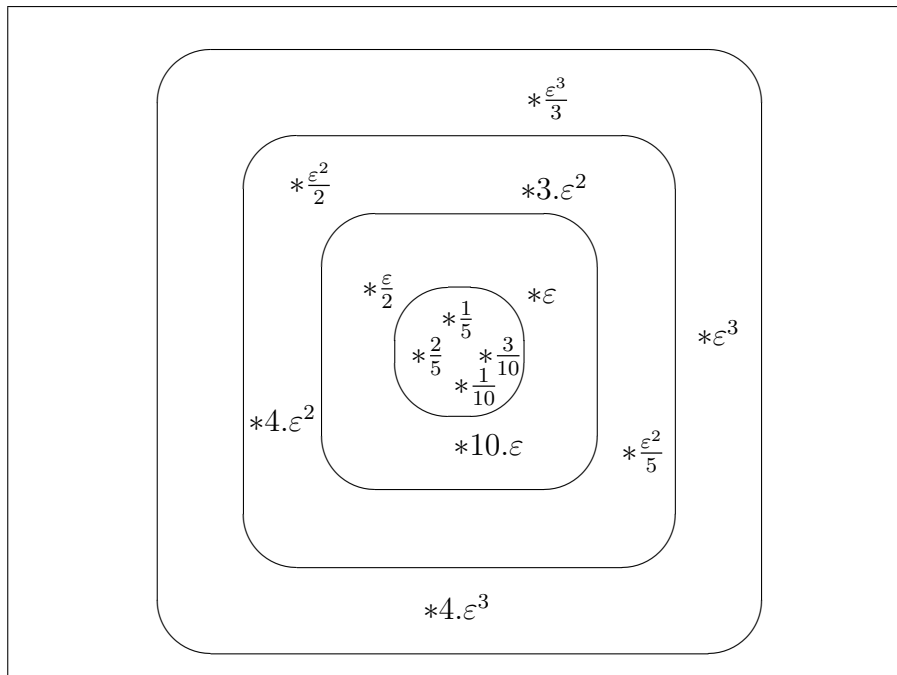


**Figure 1.** *Example of a pd-model*

### 2.1.2. *Examples.*

In this paper we will follow step by step two examples: the 'Urn' example and the 'Answering Machine' example of the introduction.

EXAMPLE 9 ('URN' EXAMPLE). — Suppose the agent is in a fair, and there is an urn containing $n = 2.k > 0$ balls which are either white or black. The agent does not know how many black balls there are in the urn but believes that it is equally probable that there are 0,1,..., or $n$ black balls in the urn. Now say there is actually 0 black balls in the urn. This situation is depicted in Fig.2. The worlds are within squares and the double bordered world is the actual world. The numbers within squares stand for the probabilities of the worlds and the propositional letter $p_i$ stands for: "there are $i$ black balls in the urn".
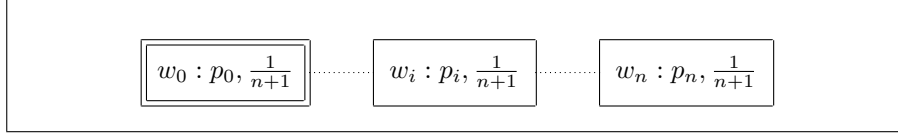
**Figure 2.** *'Urn' Example*

□

EXAMPLE 10 ('ANSWERING MACHINE' EXAMPLE). — Assume a professor (the agent) comes back after lunch to her office. She ate with a colleague who just told her at lunch about his new timetable for this year. However she does not remember quite precisely what he said, in particular she is a bit uncertain whether his 1.5 hour lecture at 2.00 pm is on Tuesday or on Thursday, and she believes 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 represent this situation in the model of Fig. 3. As in Fig. 2, the double bordered world is the actual world. The numbers within squares stand for the probabilities of the worlds and $p$ (resp. $\neg p$) stands for "her colleague has his 1.5 hour lecture on Tuesday (resp. Thursday)"[2]. □
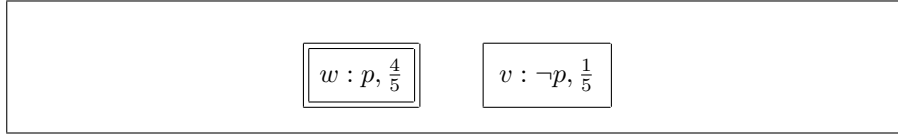


**Figure 3.** *'Answering Machine' Example*

### 2.1.3. *Static language.*

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

DEFINITION 11. — *The syntax of the language $\mathcal{L}_{St}$ is defined by*

$\phi := \perp \mid p \mid \neg\phi \mid \phi \wedge \psi \mid P(\phi) \geq x \mid C\phi$ *where* $x \in [0;1[$, *and* $p$ *ranges over propositional letters.*

*Its semantics is inductively defined by*

$M, w \models p$ *iff* $w \in V(p)$;

$M, w \models \neg\phi$ *iff not* $M, w \models \phi$

$M, w \models \phi \wedge \psi$ *iff* $M, w \models \phi$ *and* $M, w \models \psi$

---

2. Note that strictly speaking we would need two propositional variables $p$ and $p'$, where $p$ (resp. $p'$) would stand for "her colleague has his lecture on Tuesday (resp. Thursday)" because the fact that he does not have his lecture on Tuesday ($\neg p$) should not imply logically that he has his lecture on Thursday ($p'$).

$M, w \models P(\phi) \geq x$ *iff* $\sum\{P(v); M, v \models \phi\} \geq x;$

$M, w \models C\phi$ *iff* $\sum\{P(v); M, v \models \phi\} = 1.$

$M, w \models P(\phi) \geq x$ should be read "$\phi$ has a probability greater than $x$ for the agent". (Note that the world $w$ does not really play a role here.) $M, w \models C\phi$ should be read "the agent is convinced (sure) of $\phi$". We use the term conviction instead of belief because the term "belief" refers in natural language to different concepts (that we distinguish here through $\mathcal{L}_{St}$). Indeed, assume that you conjecture an arithmetical theorem $\phi$ 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 $P(\phi) \geq a$ for $a$ real and smaller than 1; the bigger $a$ is the more you "believe" in $\phi$. 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 a conviction and corresponds here to the formula $C\phi$. 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 Fermat's theorem). 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 Lenzen's in [LEN 78]. Moreover, if we define the operator $B$ by $M, w \models B\phi$ iff $M, w \models P(\phi) > 0.5$, then $B$ corresponds to Lenzen's notion of "weak belief" and $C$ and $B$ satisfy Lenzen's axioms defined in [LEN 78]. This notion of conviction also corresponds to Gärdenfors' notion of accepted formula (in a belief set) [GäR 88]. This operator $C$ satisfies the axioms K, D, 4, 5 but not the axiom T. So it does not correspond to the notion of knowledge and we do not deal with this notion in this paper. For a more in depth account on the different significations of the term "belief", see for example [LEN 78].

Moreover, we can also express in this language what would surprise the agent, and how much. Indeed, in case $x$ is an infinitesimal $\varepsilon$, $P(\phi) = \varepsilon$ should be read "the agent would be surprised with degree $\varepsilon$ if she learnt that $\phi$ ".[3] ($P(\phi) = \varepsilon$ is defined by $P(\phi) \geq \varepsilon \wedge P(\neg\phi) \geq 1 - \varepsilon$.) 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 Spohn's spirit. Indeed, $P(\neg\phi) = \varepsilon > \varepsilon' = P(\neg\phi')$ would then mean that $\phi'$ is believed more firmly than $\phi$.

Note that above, if $x$ is real then only the conceived worlds have to be considered in the sum $\sum\{P(v); M, v \models \phi\}$ because the sum of any real (different from 0) with an infinitesimal is equal to this real (see Sect. 1). Likewise, the semantics of $C$ amounts to say that $\phi$ is true at least in all the *conceived* worlds of $W$. So it is quite possible to have a surprising actual world where $\neg\phi$ is true and still the agent being convinced of $\phi$ (i.e. $C\phi$): just take $\phi$:="All swans are white" for example.

---

3. 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 real probability and still be surprised to hear them claimed by somebody else (Gerbrandy, private communication).

## 2.2. *The dynamic part*

### 2.2.1. *The notion of generic action model*

DEFINITION 12. — *A generic action model is a structure* $\mathbf{\Sigma} = (\Sigma, S, P, \{P^\Gamma; \Gamma \text{ is a}$ *maximal consistent subset of S }*$, \{Pre_\sigma; \sigma \in \Sigma\}, \sigma_0)$ *where*

*1)* $\Sigma$ *is a finite set of possible actions;*

*2)* $\sigma_0$ *is the actual action;*

*3)* $S$ *is a set of formulas of* $\mathcal{L}_{St}$ *closed under negation;*

*4)* $P^\Gamma$ *is a probabilistic measure defined on* $\mathcal{P}(\Sigma)$ *indexed by each maximal consistent subset* $\Gamma$ *of S, and assigning to each possible action* $\sigma$ *a* real *number in* $[0; 1]$*;*

*5)* $P$ *is a probabilistic measure defined on* $\mathcal{P}(\Sigma)$ *and assigning to each possible action* $\sigma \in \Sigma$ *a number in* $]0; 1]$*;*

*6)* $Pre_\sigma$ *is a function which takes as argument any propositional letter p and yields a formula of* $\mathcal{L}_{St}$*.*

*Intuitive interpretation.* Items 1,2 are similar to Definition 8. It remains to give an interpretation to items 3-6 of the definition.

*Items 3 and 4.* $S$ corresponds to the set of facts about the world that are relevant to determine the probabilities $P^\Gamma$ of Item 4. The maximal consistent sets $\Gamma$ then cover all the relevant eventualities needed for the modelling. The choice of the formulas of $S$ is left to the modeler but they should be as elementary and essential as possible in order to give rise to all the relevant eventualities. In that respect, one should avoid infinite sets $S$.

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

This probability value is real and cannot be infinitesimal (unlike $P$), and (still unlike $P$) we can have $P^\Gamma(\sigma) = 0$. This last case intuitively means that the action $\sigma$ cannot *physically* be performed in a world where $\Gamma$ is true. These operators $P^\Gamma$ generalize the binary notion of precondition in [BAL 04], [AUC 04] and [BEN 03].

REMARK 13. — In our definition of generic action model, instead of referring directly to worlds $w$ of a particular pd-model, we refer to maximal consistent subsets $\Gamma$ of a set $S$. This is for several reasons. Firstly, the determination of the probability $P^w(\sigma)$ does not depend on all the information provided by the world $w$ but often on just a part of it expressed by $\Gamma$. Secondly, for philosophical reasons, the *same* action

might be performed in different situations, modelled by different pd-models, and be perceived differently by the agent in each of these situations, depending on what her epistemic state is (see Example 18). The use of maximal consistent sets enables us to capture this *generic* aspect of actions: in our very definition we do not refer to a particular pd-model. Finally, for computational reasons, we do not want that in practice the definition of generic action models depends on a particular pd-model because we want to be able to iterate the action without having to specify at each step of the iteration the new action model related to the new pd-model. The use of maximal consistent sets allows us to do so. However, from now on and for better readability we note $P^w(\sigma) := P^{\Gamma}(\sigma)$ **for the unique $\Gamma$ such that $M, w \models \Gamma$** (such a $\Gamma$ exists because $S$ contains the negation of each formula it contains).    □

***Item 5***. $P(\sigma)$ is the probability for the agent that $\sigma$ actually occurs among $\Sigma$, determined solely by the agent's *perception* and *observation* of the action happening. This probability is independent of the agent's beliefs of the static situation which could alter and modify this determination consciously or unconsciously (see the 'answering machine' example). This probability is thus independent of the probability that the agent would have expected $\sigma$ to happen, because to determine this last probability the agent 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 actions that the agent conceives as possible candidates while one of the actions of $\Sigma$ actually takes place. The latter are actions that the agent would be surprised to learn that they actually took place while one of the other actions took place. To take up the poker example of Sect. 2.1.1, 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 $\varepsilon$) whereas the action where nothing particular happens is a conceived action (of value 1). Just as in the static case, the relative strength of the actions (conceived and surprising) is expressed by the value of the operator $P$.

***Item 6***. The function $Pre_{\sigma}$ deals with the problem of determining what facts will be true in a world after the action $\sigma$ takes place. Intuitively, $Pre_{\sigma}(p)$ represents the necessary and sufficient $Pre$condition in any world $w$ for $p$ to be true after the performance of $\sigma$ in this world $w$.

### 2.2.2. *Examples*

EXAMPLE 14 ('URN' EXAMPLE). — Consider the action whereby someone else draws a ball from the urn (which is actually a white ball) and puts it in his pocket, the agent sees him doing that but she cannot see the ball. This action is depicted in Fig. 4. Action $\sigma$ (resp. $\tau$) stands for "someone else draws a black (resp. white) ball and puts it in his pocket". The numbers within the squares stand for the probabilities $P(\sigma)$ of the possible actions. The maximal consistent sets are represented by their 'positive' components, so $\{p_i\}$ refers to the set $\{p_i, \neg p_k; k \neq i\}$.

$$\boxed{\sigma, \tfrac{1}{2}} \quad \boxed{\boxed{\tau, \tfrac{1}{2}}}$$

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

**Figure 4.** *Someone else draws a (white) ball and puts it in his pocket and the agent is uncertain whether he draws a white ($\sigma$) or a black ball ($\tau$).*

The observation and perception of the action in itself does not provide the agent any reason to have a preference between him drawing a black ball or a white ball; so we set $P(\sigma) = P(\tau) = \tfrac{1}{2}$.[4] However if the agent assumed she was in a world where there are $i$ black balls then the probability that she would (have) expect(ed) him drawing a black (resp. white) ball would be $\tfrac{i}{n}$ (resp. $1 - \tfrac{i}{n}$); so we set $P^{\{p_i\}}(\sigma) = \tfrac{i}{n}$ and $P^{\{p_i\}}(\tau) = 1 - \tfrac{i}{n}$. Moreover there cannot be $n$ black balls in the urn after he put one ball in his pocket; so we set $Pre_\sigma(p_n) = \perp$ and $Pre_\tau(p_n) = \perp$. But if he draws a black ball then there is one black ball less; so we set $Pre_\sigma(p_i) = p_{i+1}$ for all $i < n$. Otherwise if he draws a white ball the number of black balls remains the same; so we set $Pre_\tau(p_i) = p_i$ for all $i < n$. $\square$

EXAMPLE 15 ('ANSWERING MACHINE' EXAMPLE). — Assume now that when the professor enters her office, she finds a message on her answering machine from her colleague. He tells her that he will bring her back a book he had borrowed 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 she cannot distinguish precisely whether he said Tuesday or Thursday. Nevertheless she considers more probable of having heard Tuesday rather than Thursday.

We model this action in Fig. 5. $\sigma$ stands for "her colleague says that he will come on Tuesday between 2.00 pm and 4.00 pm" and $\tau$ stands for "her colleague says that he will come on Thursday between 2.00 pm and 4.00 pm". $P(\sigma)$ and $P(\tau)$ represent the probabilities the agent assigns to $\sigma$ and $\tau$ on the sole basis of what she has heard and distinguished from the answering machine (and are depicted within the squares). These probabilities are determined on the basis of her sole perception of the message. On the other hand, $P^{\{p\}}(\sigma)$ is the probability that she would have expected her colleague to say that he will come on Tuesday (rather than Thursday) if she assumed that his lecture was on Tuesday. This probability can be determined objectively. Indeed, because we assume that her colleague has a lecture on Tuesday from 2.00 pm to 3.30

---

4. We could nevertheless imagine some exotic situation where the agent's observation of him drawing the ball would give her some information on the color of the ball he is drawing. For example, if black balls were much heavier than white balls and the agent sees him having difficulty drawing a ball, she could consider more probable that he is drawing a black ball.

$$\boxed{\sigma, \tfrac{3}{5}} \qquad \boxed{\tau, \tfrac{2}{5}}$$

$$S = \{p, \neg p\}.$$
$$P^{\{p\}}(\sigma) = \tfrac{1}{5},\ P^{\{p\}}(\tau) = \tfrac{4}{5}$$
$$P^{\{\neg p\}}(\sigma) = \tfrac{4}{5},\ P^{\{\neg p\}}(\tau) = \tfrac{1}{5}.$$
$$Pre_\sigma(p) = p, \text{ and } Pre_\tau(p) = p$$

**Figure 5.** *The agent is uncertain whether her colleague says that he will come on Tuesday ($\sigma$) or on Thursday ($\tau$) between 2.00 pm and 4.00 pm, but she considers more probable having heard Tuesday than Thursday.*

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 $P^{\{p\}}(\sigma) = \frac{0.5hr}{2.5hrs} = \frac{1}{5}$. Similarly, $P^{\{p\}}(\tau) = \frac{2hrs}{2.5hrs} = \frac{4}{5}$. Finally, the message does not change the fact of her colleague having a lecture or not on Tuesday; so we set $Pre_\sigma(p) = p$ and $Pre_\tau(p) = p$.

$\square$

### 2.3. *The update mechanism*

2.3.1. *The notion of update product*

DEFINITION 16. —

Given a pd-model $M = (W, P, V, w_0)$ and a generic action model $\Sigma = (\Sigma, S, P, \{P^\Gamma; \Gamma$ is a m.c. subset of $S\}, \{Pre_\sigma; \sigma \in \Sigma\}, \sigma_0)$, we define their update product to be the pd-model $M \otimes \Sigma = (W', P', V', w_0')$, where:

1) $W' = \{(w, \sigma) \in W \times \Sigma; P^w(\sigma) > 0\}$.

2) We set

$$P'(\sigma) = \frac{P(\sigma).P^W(\sigma)}{\sum\{P(\tau).P^W(\tau); \tau \in \Sigma\}} \text{ where } P^W(\sigma) = \sum\{P(v).P^v(\sigma); v \in W\}.$$

Then

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

3) $V'(p) = \{(w, \sigma) \in W'; M, w \models Pre_\sigma(p)\}$.

4) $w_0' = (w_0, \sigma_0)$.

***Intuitive interpretation and motivations.***

***Items 1 and 4*** As in BMS, in the new model we consider all the possible worlds $(w, \sigma)$ resulting from the performance of the possible action $\sigma$ in the possible world

$w$, granted that this action $\sigma$ can physically take place in $w$ (i.e. $P^w(\sigma) > 0$). The new actual world is the result of the performance of the actual action $\sigma_0$ in the actual world $w_0$.

REMARK 17. — When it comes to modelling epistemic states of agents, there are two (exclusive) approaches to follow: the objective approach and the subjective approach. In the first approach, the modeler has an objective, external and omniscient knowledge of the agent's epistemic state and the models she builds to represent the agent's epistemic state are thus perfect and correct. This approach has applications in game theory (and economics) where an objective point of view is necessary to explain past economical behaviors for example. In the second approach, the modeler is the agent at stake herself and as such her perception of the situation might be erroneous because she is part of the situation. So the models she builds to represent her own epistemic state might be erroneous. This approach has applications in artificial intelligence where agents/robots need to have a formal representation of the surrounding world. These two approaches lead to two different kinds of formalism and ways to formalize belief change.

In this paper we follow the objective approach. With this assumption, our pd-models and generic action models are correct. So the actual world and actual action of our models do correspond to the actual world and actual action in reality. It is then natural to assume that the actual action $\sigma_0$ can physically be performed in the actual world $w_0$: $P^{w_0}(\sigma_0) > 0$. Hence, the existence of the (actual) world $(w_0, \sigma_0)$ is justified. $\qquad\square$

***Item 2***. We want to determine $P'(w, \sigma) = P(W \cap A)$, where $W$ stands for the event 'we were in world $w$ before $\sigma$ occurred' and $A$ for 'action $\sigma$ just occurred'. More formally, in the probability space $W'$, $W$ stands for $\{(w, \tau); \tau \in \Sigma\}$ and $A$ for $\{(v, \sigma); v \in W\}$ 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 $\Sigma$.
Probability theory tells us that

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

We first determine $P(W|A)$, i.e. the probability that the agent was in world $w$ *given* the extra assumption that action $\sigma$ occurred in this world. We reasonably claim

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

That is to say, we *conditionalize* the probability of $w$ for the agent (i.e. $P(w)$) to the worlds where the action $\sigma$ took place and that may correspond for the agent to the actual world $w$ (i.e. $\{v; v \in W \text{ and } P^v(\sigma) > 0\}$). 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 $\sigma$ occurred in $w$. So the worlds where the action $\sigma$ 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 $P(A)$, which we also denote $P'(\sigma)$; that is to say the probability for the agent that $\sigma$ occurred. We claim that

$$P(A) = P_1.P_2$$

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

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

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

Finally, we normalize $P'(\sigma)$ on the set of actions $\Sigma$ to get a probabilistic space. We get

$$P(A) = P'(\sigma) = \frac{P(\sigma).P^W(\sigma)}{\sum\{P(\tau).P^W(\tau); \tau \in \Sigma\}} \text{ where } P^W(\sigma) = \sum\{P(v).P^v(\sigma); v \in W\}.$$

We can easily check that $\sum\{P'(w,\sigma); (w,\sigma) \in W'\} = 1$, which ensures that $P'$ is a probability measure on $\mathcal{P}(W')$.

***Item 3***. Intuitively, this formula says that a fact $p$ is true after the performance of $\sigma$ in $w$ iff the necessary and sufficient precondition for $p$ to be true after $\sigma$ was satisfied in $w$ before the action occurred.[5]

───────────────

5. In [AUC 05], the solution of determining which propositional facts were true after an update was very similar to Reiter's solution to the frame problem [REI 01]. It turns out that the solution we propose here is equivalent to the previous one and was first proposed by Renardel De Lavalette [REN 04] and Kooi [KOO 05], [DIT 05]. Indeed, in [AUC 05] 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 $\sigma$ was $Post_\sigma^+(p) \vee (p \wedge Post_\sigma^-(p))$ which is shortened here to $Pre_\sigma(p)$.

2.3.2. *Examples*

EXAMPLE 18 ('URN' EXAMPLE). —   Assume now that someone else just drew a (white) ball from the urn and put it in his pocket, action depicted in Fig.4. Then because the agent considered equally probable that there was 0,1,..., or $n$ black balls in the urn, she would expect that the other person drew a black ball or a white ball with equal probability. That is indeed the case:

$$P^W(\sigma) = \sum\{P(v).P^v(\sigma); v \in W\} = \sum\{\tfrac{1}{n+1}.\tfrac{i}{n}; i = 0..n\} = \tfrac{1}{2} = P^W(\tau).$$

Independently from that, her perception of the action did not provide her any reason to prefer $\sigma$ over $\tau$ (i.e. $P(\sigma) = P(\tau)$). So, in the end she should believe equally that the other agent drew a black ball or a white ball, and this is indeed the case: $P'(\sigma) = P'(\tau)$.

If we perform the full update mechanism, then we get the pd-model of Fig.6. In this model all the worlds are equally probable for the agent. Note that there cannot be $n$ black balls in the urn $(p_n)$ any longer since one of them has been withdrawn.

Now consider another scenario where this time the agent initially believes that there are more black balls than white balls (for example she believed somebody else that told her so in the beginning). This can be modelled by assigning the initial probabilities $P(w_i) = \varepsilon$ for $i = 0, .., k$ and $P(w_i) = \tfrac{1}{k}$ 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 somebody else draws a ball from the urn and we compute again the probabilities of the actions $\sigma$ and $\tau$, we get what we expect, namely that the agent considers more probable that a black ball has been withdrawn rather than a white ball:

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

$\square$



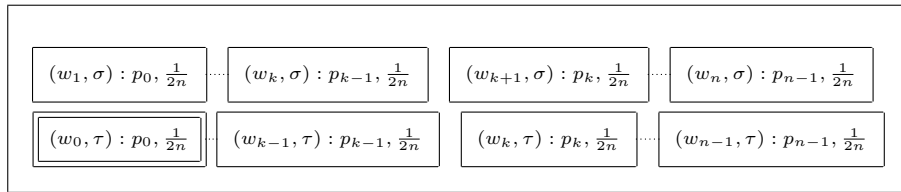| $(w_1, \sigma) : p_0, \tfrac{1}{2n}$ | $(w_k, \sigma) : p_{k-1}, \tfrac{1}{2n}$ | $(w_{k+1}, \sigma) : p_k, \tfrac{1}{2n}$ | $(w_n, \sigma) : p_{n-1}, \tfrac{1}{2n}$ |
| $(w_0, \tau) : p_0, \tfrac{1}{2n}$ | $(w_{k-1}, \tau) : p_{k-1}, \tfrac{1}{2n}$ | $(w_k, \tau) : p_k, \tfrac{1}{2n}$ | $(w_{n-1}, \tau) : p_{n-1}, \tfrac{1}{2n}$ |

**Figure 6.** *Situation of the urn example after somebody else drew a white ball and put it in his pocket ($\sigma$).*

EXAMPLE 19 ('ANSWERING MACHINE' EXAMPLE). —   Now that the professor has heard the message, she updates her 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 she computes her action probabilities $P'(\sigma)$ and $P'(\tau)$.

After this computation, the probability $P'(\sigma)$ that her colleague said that he will come on Tuesday is a combination of: (1) how much she would have expected him to say so, based on what she knew and believed of the situation, and (2) what she actually distinguished and heard from the answering machine.

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

Now consider a second scenario where this time she initially believes with a higher probability that he has got a lecture on Thursday than on Tuesday ($P(w) = \frac{1}{5}$ and $P(v) = \frac{4}{5}$ in Fig.2). This time her belief that she heard him saying that he will come on Tuesday (resp. Thursday) is strengthened (resp. weakened) by her independent expectation of him to say so: $P'(\sigma) > P(\sigma) > P(\tau) > P'(\tau)$.

$\square$

EXAMPLE 20 ('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 initial situation, except that her beliefs are different. This time she is convinced that her colleague's lecture is on Thursday and she is also convinced that she heard him saying that he would come on Thursday. Formally, everything remains the same except that now in Fig. 3 $P(w) = \varepsilon$ and $P(v) = 1$, and in Fig. 5 $P(\sigma) = \varepsilon$ and $P(\tau) = 1$. If we apply the full update mechanism with these new parameters we get the model depicted in Fig. 7. In this model she is convinced that he has got a lecture on Thursday and that he said he would come on Thursday (world $(v, \tau)$). Moreover, she would be surprised (with degree $5\varepsilon$) if she learnt that her colleague has got his lecture on Tuesday *or* he said he will come on Tuesday (worlds $(w, \tau)$, $(v, \sigma)$ and $(w, \sigma)$; remember that $5\varepsilon + 4\varepsilon^2 = 5\varepsilon$). But she would be much more surprised (with degree $\varepsilon^2$) if she learnt that he has got his lecture on Tuesday *and* he said he will come on Tuesday (world $(w, \sigma)$), because that contradicts twice her original convictions.

Now later her colleague tells her that he has got his lecture on Tuesday, then she will have to revise her beliefs. The action model of this public announcement is depicted in Fig. 8. The resulting model is depicted in Fig. 9. In this model, she now believes that he has got his lecture on Tuesday, but she is still convinced that her colleague said that he will come on Thursday because no new information has contradicted this. This is made possible thanks to the (global) ranking of surprising worlds
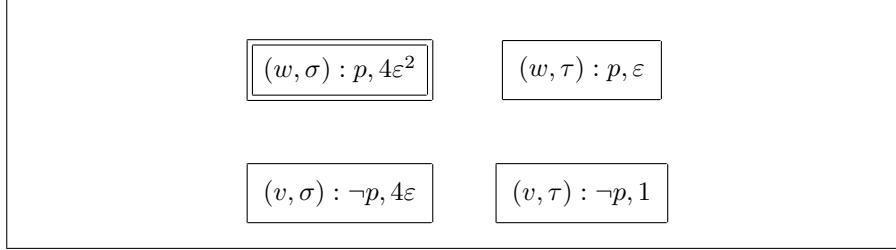
$$\boxed{\boxed{(w,\sigma) : p, 4\varepsilon^2}} \qquad \boxed{(w,\tau) : p, \varepsilon}$$

$$\boxed{(v,\sigma) : \neg p, 4\varepsilon} \qquad \boxed{(v,\tau) : \neg p, 1}$$

**Figure 7.** *Situation after she believed that her colleague's lecture is on Thursday ($\neg p$) and she heard him saying that he would come on Thursday ($\tau$).*

by infinitesimals. What happened during this revision process is that the least surprising world where $p$ is true became the only conceived world.
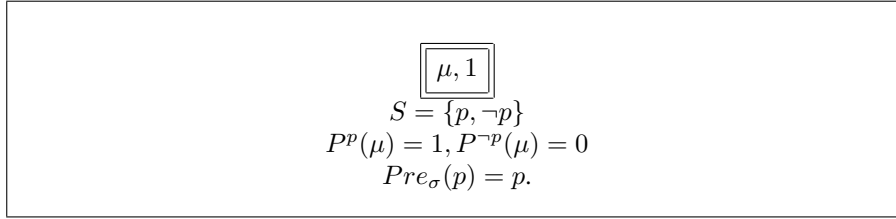
□

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

**Figure 8.** *Her colleague announces to her that his lecture is on Tuesday.*

$$\boxed{\boxed{((w,\sigma),\mu) : p, 4\varepsilon}} \qquad \boxed{((w,\tau),\mu) : p, 1}$$
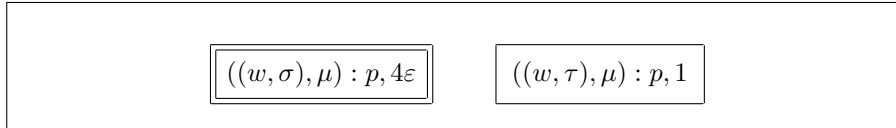
**Figure 9.** *Situation after her colleague announced his lecture is on Tuesday ($\mu$) and she then revised her beliefs.*

## 3. Extension to the multi-agent case.

In this section we only sketch a way to extend this framework to the multi-agent case. At the address *ftp://ftp.irit.fr/IRIT/LILAC/Aucher_jancl.pdf* can be found the detailed formalism. In the sequel, $G$ is a fixed set of agents[6].

---

6. Another possible and equivalent extension consists in simply indexing the probabilities $P^\Gamma$ of the generic action model by the agents $j \in G$; and indexing the probabilities $P$ of the pd-

We introduce a set of equivalence relations $\sim_j$ on the set of worlds $W$, modelling the rough uncertainty of each agent $j \in G$. Then each $\sim_j$ equivalence class will be considered as an instance of a "single-agent" pd-model with probability measure $P_j$ and should then fulfill the same conditions.

We do similarly for the generic action models by introducing a set of equivalence relations $\sim_j$ on the set of possible actions $\Sigma$ indexed by the agents $j \in G$. Each $\sim_j$ equivalence class will be considered as an instance of a "single-agent" generic action model with probability measures $P_j$ and $P_j^\Gamma$ and should fulfill the same conditions. Moreover, we assume that the fact that an action cannot physically take place in a world is a public fact among agents. So we should add in the definition of the generic action model that for each possible action $\sigma$ and agent $j_0$, if $P_{j_0}^\Gamma(\sigma) = 0$ then $P_j^\Gamma(\sigma) = 0$ for all $j \in G$.

Finally, the update mechanism is the same as for the "single-agent" case: the $P$ operators just have to be indexed by the agents $j \in G$. However we also have to deal with the new component $\sim_j$. As in BMS, we set $(w, \sigma) \sim'_j (v, \tau)$ iff $w \sim_j v$ and $\sigma \sim_j \tau$ because the uncertainty relations $\sim_j$ for the pd-model and the generic action model are independent from one another.

## 4. 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 [LEH 96], lexicographic probability [BLU 91], big-stepped probabilities [DUB 04] and Weydert's cumulative algebra [WEY 94]. All these proposals are very similar in spirit and seem to be equivalent in one way or another. We showed in Theorem 7 that our structure $\mathbb{V}$ is actually an extension of a structure $\mathbb{V}'$ isomorphic to a cumulative algebra. So Weydert's comparisons with Spohn's theory [SPO 88] and possibility theories [DUB 91] transfer. In particular the ranking $\mathbb{V}^0$ determines a *global* ranking of worlds similar in spirit to Spohn's degrees

--------

model (resp. generic action model) by the agents $j \in G$ and possible worlds $w$ (resp. possible actions $\sigma$) but allowing at the same time these probability measures to take the value 0 (note that it was not the case for the single-agent case). $P_{j,w}(v) = 0$ would then mean that if the agent $j$ is in world $w$ then the world $v$ is not relevant to describe her own epistemic state (and similarly for possible actions). However, if we want to have a framework equivalent to the first one, we have to add constraints to the $P_{j,w}$ (and $P_{j,\sigma}$), namely

1) $P_{j,w}(w) > 0$ for all $w \in W$,

2) if $P_{j,w}(v) > 0$ then for all $u \in W$ $P_{j,w}(u) = P_{j,v}(u)$,

3) $\sum \{P_{j,w}(v); v \in W\} = 1$.

Note that these constraints look similar to seriality (3), transitivity and euclidianity (2) constraints. Finally, the update product is the same except that the probabilities $P$ have to be replaced by $P_{j,w}$ (or $P_{j,\sigma}$) and the probabilities $P^\Gamma$ by $P_j^\Gamma$.

of disbelief or possibility degrees in possibility theory[7]. (The global degree of a world $w$ can be defined by $\alpha((P(w))^0) \in P(w)^0$ where $\alpha$ is a choice function; see the appendix.) 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 correspond to Spohn's plausibility degrees (although the order has to be reversed). Likewise with possibility theory. In Fig. 1 of Sect. 2.1, the worlds with the same global degree are located between two consecutive spheres. But among the worlds of the same global degree exists also a *local* ranking corresponding to the usual order relation. So for example, in the picture we have the local ranking $4.\varepsilon^2 > 3.\varepsilon^2$, 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 21. — In the literature (including myself in [AUC 04]), 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 [GäR 88]. □

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

THEOREM 22. — *If, as in the AGM theory, we restrict our attention to propositional beliefs, then in the case of a public announcement our update mechanism satisfies the eight AGM postulates for belief revision.*

Suppose $\phi$ is false in every conceived world, when we revise by $\phi$ then the surprising worlds where $\phi$ 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 probability 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 [KOO 03] is based on the static approach by Fagin and Halpern in [FAG 94]. Unlike us, probability is not meant only to model the agents' epistemic states. In that respect, his

---

7. $\mathbb{V}^0$ is defined as being the quotient structure of $\mathbb{V}'$ by the equivalence relation $\approx^0$, which is itself defined by $x \approx^0 y$ iff $\frac{x}{y}$ is real different from 0 if $y \neq 0$ and $x = 0$ if $y = 0$ (see the appendix).

probability measures are defined relatively to each world without any constraint on them. Moreover he only deals with public announcement. But 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 zero 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 different from zero. Finally, he does not consider actions changing facts (he tackles this topic independently in [KOO 05]).

**Comparison with van Benthem's system.** van Benthem's early system [BEN 03] is similar to ours in its spirit and goals. However he does not introduce the probabilities $P^W(\sigma)$ and $P(\sigma)$ but only a single $P^w(\sigma)$. Hence, his probabilistic update rule is different. The intended interpretation of his $P^w(\sigma)$ seems also to be different from ours if we refer to his example. Anyway, his discussion and comparison with the Bayesian setting in his Sect. 5 are still valid here.

A more elaborated version of his system which is very similar to ours has been developed independently by him, Kooi and Gerbrandy [BEN 06]. In this system they do have a second probability measure $P$ for the action models whose intended meaning is the same as ours. Nevertheless, their probabilistic update rule is still different and does not comply to the Jeffrey update, contrary to ours. They also study some parameterized versions of their probabilistic rule (which could be done here too) and they show that one of them actually complies to the Jeffrey update. They provide a sketch of a completeness proof via reduction axioms. However, they do not resort to infinitesimals to represent epistemic states and thus can neither express degrees of potential surprise nor allow for belief revision. Finally they do not consider actions changing facts.

**Comparison with the situation calculus of Bacchus, Halpern and Levesque.** Their system [BAC 99] can be viewed as the counterpart of van Benthem's early 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 normalization). So what applies to van Benthem's early system applies here too. 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 [BOU 98] deals with noisy observations. Their approach is semantically driven like ours. However they use a different framework called observation system based on the notion of (ranked) interpreted system. 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 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 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 for example express, as

they do, that the observation is "credible". 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.

## 5. Conclusion

In order to represent with most accuracy the agent's 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 infinitesimals (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 happening and our expectation of it to happen, we have proposed a way to model this phenomenon. Incidentally, 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 may change the 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. However this system can be of use as it is in several areas. Firstly, in game theory where the kinds of phenomena we studied are quite current. Secondly, in psychology if we want to devise realistic formal models of belief change. Finally, the logical dynamics we modelled could be used in some way or another in artificial intelligence since they are the hallmark of rational and efficient reasoners.

Seminar "belief change in rational agents: perspectives from artificial intelligence, philosophy, and economics" for their attention, questions and comments.

## 6. References

[ADA 75]  ADAMS E. W., *The Logic of Conditionals*, vol. 86 of *Synthese Library*, Springer, 1975.

[AUC 04]  AUCHER G., "A Combined System for Update Logic and Belief Revision", BAR-LEY M., KASABOV N. K., Eds., *Intelligent Agents and Multi-Agent Systems, 7th Pacific Rim International Workshop on Multi-Agents (PRIMA 2004)*, vol. 3371 of *LNCS*, Springer, 2004, p. 1–17, Revised Selected Papers.

[AUC 05]  AUCHER G., "How Our Beliefs Contribute to Interpret Actions.", PECHOUCEK M., PETTA P., VARGA L. Z., Eds., *Multi-Agent Systems and Applications IV, 4th International Central and Eastern European Conference on Multi-Agent Systems (CEEMAS 2005)*, vol. 3690 of *LNCS*, Springer, 2005, p. 276–285.

[BAC 99]  BACCHUS F., HALPERN J., LEVESQUE H., "Reasoning about Noisy Sensors and Effectors in the Situation Calculus", *Artificial Intelligence*, vol. 111, num. 1-2, 1999, p. 171–208.

[BAL 98]  BALTAG A., MOSS L., SOLECKI S., "The logic of Common Knowledge, Public Announcement, and Private Suspicions", GILBOA I., Ed., *Proceedings of the 7th conference on theoretical aspects of rationality and knowledge (TARK98)*, 1998, p. 43–56.

[BAL 04]  BALTAG A., MOSS L., "Logic for Epistemic Program", *Synthese*, vol. 139, num. 2, 2004, p. 165–224.

[BEN 03]  VAN BENTHEM J., "Conditional Probability Meets Update Logic.", *Journal of Logic, Language and Information*, vol. 12, num. 4, 2003, p. 409–421.

[BEN 06]  VAN BENTHEM J., GERBRANDY J., KOOI B., "Dynamic Update with Probability", report , march 2006, ILLC.

[BLU 91]  BLUME L., BRANDENBURGER A., DEKEL E., "Lexicographic Probabilities and Choice under Uncertainty", *Econometrica*, vol. 59, num. 1, 1991, p. 61–79.

[BOU 98]  BOUTILIER C., HALPERN J., FRIEDMAN N., "Belief Revision with Unreliable Observations", *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI 1998)*, 1998, p. 127–134.

[DIT 05]  VAN DITMARSCH H. P., VAN DER HOEK W., KOOI B. P., "Dynamic epistemic logic with assignment.", DIGNUM F., DIGNUM V., KOENIG S., KRAUS S., SINGH M. P., WOOLDRIDGE M., Eds., *AAMAS*, ACM, 2005, p. 141–148.

[DUB 91]  DUBOIS D., PRADE H., "Possibilistic Logic, Preferential Model and Related Issue", *Proceedings of the 12th International Conference on Artificial Intelligence (IJCAI)*, Morgan Kaufman, 1991, p. 419–425.

[DUB 04]  DUBOIS D., FARGIER H., "A Unified Framework for Order-of-Magnitude Confidence Relation", *Twentieth Conference in Artificial Intelligence*, 2004, p. 138–145.

[FAG 94]  FAGIN R., HALPERN J., "Reasoning about Knowledge and Probability", *Journal of the ACM (JACM)*, vol. 41, num. 2, 1994, p. 340–367.

[GäR 88]  Gärdenfors P., *Knowledge in Flux (Modeling the Dynamics of Epistemic States)*, Bradford/MIT Press, Cambridge, Massachusetts, 1988.

[KEI 86]  Keisler H. J., *Elementary Calculus:  an Approach Using Infinitesimals*,  Prindle and Weber & Schmidt, 1986,  Online edition on the website http://www.math.wisc.edu/ keisler/calc.html.

[KOO 03]  Kooi B., "Probabilistic Dynamic Epistemic Logic", *Jourbal of Logic, Language and Information*, vol. 12, num. 4, 2003, p. 381–408.

[KOO 05]  Kooi B., "As the World Turns: On the Logic of Public Update", Manuscript, 2005.

[LEH 96]  Lehmann D., "Generalized Qualitative Probability: Savage Revisited", Horvitz E., Jensen F., Eds., *Twelfth Conference on Uncertainty in Artificial Intelligence*, Portland, Oregon, August 1996, Morgan Kaufmann, p. 381–388.

[LEN 78]  Lenzen W., *Recent Work in Epistemic Logic*, Acta Philosophica 30, North Holland Publishing Company, 1978.

[REI 01]  Reiter R., *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. 251–275.

[SPO 88]  Spohn W., "A General Non-Probability Theory of Inductive Reasoning", Schachter R., Levitt T., Kanal L., Lemmer J., Eds., *Uncertainty in Artificial Intelligence 4*, North-Holland, 1988, p. 149–158.

[WEY 94]  Weydert E., "General Belief Measure",  de Mántaras R. L., Poole D., Eds., *Tenth Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann, 1994, p. 575–582.

## A.  Proof of Theorem 3., Lemma 5. and Theorem 6.

THEOREM 23 (THEOREM 3.). — *The quotient structure $\mathbb{V} = (^*\mathbb{R}^+/_{\approx}, \overline{+}, \overline{\div})$ is a semi-field.*

PROOF. — We only need to prove that $\overline{+}$ and $\overline{\div}$ 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 $\overline{x}\,\overline{\div}\,\overline{y} = \overline{x'}\,\overline{\div}\,\overline{y'}$.

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 $x, y \geq 0$.

$\overline{x \overline{+} y} = \overline{x' \overline{+} y'}$ 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 $\overline{x \div y} = \overline{x' \div y'}$.

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 $x.y \neq 0$.

$\overline{x \div y} = \overline{x' \div y'}$ 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.

∎

LEMMA 24 (LEMMA 5.). — *If $\overline{x} \lesssim \overline{y}$ 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

$0 \lneqq \overline{y}$ 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 $x' > 0$.

$x_0 \leq 0 < x'$

$1 = \frac{x_0}{x_0} \le 0 < \frac{x'}{x_0}$

$1 \le 0 \le St(\frac{x'}{x_0}) = 1$ because $x', x_0 \in \overline{x}$

i.e. $0 = 1$ which is again counterintuitive.

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

$\blacksquare$

THEOREM 25 (THEOREM 6.). — *The structure* $(\mathbb{V}, \underset{\approx}{\lesssim})$ *is an ordered semi-field.*

PROOF. — First we prove a lemma:

LEMMA 26. — $\underset{\approx}{\lesssim}$ *is a total order on* $^*\mathbb{R}^+$ *such that*

*1) if* $\overline{x} \underset{\approx}{\lesssim} \overline{y}$ *then* $\overline{x+z} \underset{\approx}{\lesssim} \overline{y+z}$,

*2) if* $0 \underset{\approx}{\lesssim} \overline{x}$ *and* $0 \underset{\approx}{\lesssim} \overline{y}$ *then* $0 \underset{\approx}{\lesssim} \overline{x \cdot y}$.

PROOF. — Follows easily from Lemma 24 (Lemma 5) and the fact that $\le$ is a total order on $^*\mathbb{R}^+$ satisfying also conditions 1 and 2 above. $\blacksquare$

The proof then follows easily from the lemma above and Theorem 23 (Theorem 3). $\blacksquare$


## B.  Proof of Theorem 7.

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 27. — *Let* $x, y \in \mathbb{V}'$, *we set*

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

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

DEFINITION 28. — *We define the ranking* $\mathbb{V}^0 = (V^0, +^0, \cdot^0, 0^0, 1^0, \le^0)$ *by the quotient structure of* $\mathbb{V}'$ *by the equivalence relation* $\approx^0$.

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

THEOREM 29 (THEOREM 7.). — $\mathbb{V}'$ *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 $g(x^0, y) = \alpha(x^0).y$. $\blacksquare$

## C. Proof that our System Satisfies the eight AGM Postulates

To check whether the AGM postulates are fulfilled, we first need to define 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 $\phi$, depicted in Fig. 10.

$$\boxed{\boxed{\phi!}}$$

$$S = \{\phi, \neg\phi\} \text{ and } P^{\{\phi\}}(\sigma) = 1, P^{\{\neg\phi\}}(\sigma) = 0.$$

**Figure 10.** *generic action model for the action 'public announcement of the propositional formula $\phi$'*

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.

DEFINITION 30. — *Let $M = (W, P, V, w_0)$ be a pd-model. Let $\phi, \psi$ be propositional formulas. We define*

$$M, w \models [\phi!]\psi \text{ iff, if } M, w \models \phi \text{ then } M(\phi!), (w, \phi!) \models \psi,$$

*where $M(\phi!)$ is the pd-model defined by $M(\phi!) := (W' := \{(w, \phi!); M, w \models \phi\}, \{P(w, \phi!) = \frac{P(w)}{\sum\{P(v); v \in W \text{ and } M, v \models \phi\}}; (w, \phi!) \in W'\}, V' := V, w_0' := (w_0, \phi!))$*

DEFINITION 31. —

*Let $\mathcal{L}$ be the propositional language. We define*

– *the belief set $K = \{\psi \in \mathcal{L}; M, w_0 \models C\psi\}$,*
– *the revision of the belief set $K$ by $\phi$, $K * \phi = \{\psi \in \mathcal{L}; M, w_0 \models [\phi!]C\psi\}$,*
– *the expansion of the belief set $K$ by $\phi$, $K + \phi = \{\psi \in \mathcal{L}; M, w_0 \models C(\phi \to \psi)\}$.*

THEOREM 32. — *The function * defined in Definition 31 satisfies the 8 AGM postulates.*

PROOF. —

First we recall the 8 AGM postulates.

**(K*1)** For any formula $\phi$ and any belief set $K$, $K * \phi$ is a belief set. ('Closure')

**(K*2)** $\phi \in K * \phi$. ('Success')

**(K*3)** $K * \phi \subseteq K + \phi$. ('Expansion 1')

**(K\*4)** If $\neg\phi \notin K$, then $K + \phi \subseteq K * \phi$. ('Expansion 2')

**(K\*5)** $K * \phi = K_\perp$ only if $\vdash \neg\phi$. ('Consistency preservation')

**(K\*6)** If $\vdash \phi \leftrightarrow \phi'$, then $K * \phi = K * \phi'$. ('Extensionality')

**(K\*7)** $K * \phi \wedge \phi' \subseteq (K * \phi) + \phi'$. ('Conjunction 1')

**(K\*8)** If $\neg\phi' \notin K * \phi$, then $(K * \phi) + \phi' \subseteq K * \phi \wedge \phi'$. ('Conjunction 2, Rational Monotony')

Now we prove the theorem.

**(K\*1)** This postulate is clearly satisfied.

**(K\*2)** This postulate is satisfied, 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)** $M, w_0 \models [\phi!]C\psi$

$\Leftrightarrow \sum\{P(w, \phi!); M(\phi!), (w, \phi!) \models \psi\} = 1$

$\Leftrightarrow \sum\{\frac{P(w)}{\sum\{P(v); v \in W \text{ and } M, v \models \phi\}}; w \in W \text{ and } M, w \models \phi \text{ and } M, w \models \psi\} = 1$
because propositional formulas are persistent.

$\Leftrightarrow \sum\{P(w); w \in W \text{ and } M, w \models \phi \text{ and } M, w \models \psi\} = \sum\{P(w); w \in W \text{ and } M, w \models \phi\}$

$\Rightarrow^* \sum\{P(w); w \in W \text{ and } M, w \models \phi \rightarrow \psi\} = 1$ (1)

$\Leftrightarrow M, w_0 \models C(\phi \rightarrow \psi)$

**(K\*4)** $\neg\phi \notin K$

$\Leftrightarrow M, w_0 \models \neg C\neg\phi$

$\Leftrightarrow$ there is a conceived world $w \in W$ such that $M, w \models \phi$ (H).

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

Formula (1) tells us that for all conceived worlds $w \in W$ $M, w \models \phi \rightarrow \psi$. So,

$\sum\{P(w); w \in W \text{ and } M, w \models \phi\}$

$= \sum\{P(w); w \in W \text{ and } M, w \models \phi \text{ and } w \text{ conceived }\}$ by (H), because if $v$ is conceived and $v'$ surprising then $P(v) + P(v') = P(v)$ (see Sect.1)

$= \sum\{P(w); w \in W \text{ and } M, w \models \phi \text{ and } M, w \models \psi \text{ and } w \text{ conceived }\}$ by (1)

$= \sum\{P(w); w \in W \text{ and } M, w \models \phi \text{ and } M, w \models \psi\}$ because there does exist a conceived world $w$ satisfying $\phi \wedge \psi$ by (H) and (1).

**(K\*5)** It is fulfilled because $K * \phi \neq K_\perp$.

Indeed, otherwise $M, w_0 \models [\phi!]C \perp$

but $\models \neg C \perp$, so $\models [\phi!]\neg C \perp$.

Then $M, w_0 \models [\phi!] \perp$.

That is $M, w_0 \models \neg \phi$

which is wrong because by assumption $P^{w_0}(\sigma) > 0$ (see Remark 17). This means by definition of the action model that $M, w_0 \models \phi$.

**(K\*6)** This postulate is clearly satisfied.

**(K\*7)** First note that $M, w_0 \models [\phi!]C(\phi' \to \psi) \Leftrightarrow \phi \in K^*\phi + \phi'$

$M, w_0 \models [\phi \wedge \phi']C\psi$

$\Leftrightarrow$ if $M, w_0 \models \phi \wedge \phi'$ then $M(\phi \wedge \phi'!), (w_0, \phi \wedge \phi'!) \models C\psi$

$\Leftrightarrow M(\phi \wedge \phi'!), (w_0, \phi \wedge \phi'!) \models C\psi$ because $M, w_0 \models \phi \wedge \phi'$ (see Remark 17)

$\Leftrightarrow \sum\{P(w, \phi \wedge \phi'!); w \in W \text{ and } M, w \models \phi \wedge \phi' \text{ and } M, w \models \psi\} = 1$

$\Leftrightarrow \sum\{\frac{P(w)}{\sum\{P(v); v \in W \text{ and } M, v \models \psi \wedge \phi'\}}; w \in W \text{ and } M, w \models \phi \wedge \phi' \text{ and } M, w \models \psi\} = 1$

$\Leftrightarrow \sum\{P(w); w \in W \text{ and } M, w \models \phi \wedge \phi' \text{ and } M, w \models \phi\} = \sum\{P(w); w \in W \text{ and } M, w \models \phi \wedge \phi'\}$

$\Leftrightarrow \sum\{\frac{P(w)}{\sum\{P(v); v \in W \text{ and } M, v \models \phi\}}; w \in W \text{ and } M, w \models \phi \wedge \phi' \text{ and } M, w \models \psi\} = \sum\{\frac{P(w)}{\sum\{P(v); v \in W \text{ and } M, v \models \phi\}}; w \in W \text{ and } M, w \models \phi \wedge \phi'\}$

$\Leftrightarrow \sum\{P(w, \phi!); (w, \phi!) \in W' \text{ and } M, w \models \phi \wedge \phi' \text{ and } M, w \models \psi\} = \sum\{P(w, \phi!); (w, \phi!) \in W' \text{ and } M, w \models \phi \wedge \phi'\}$

$\Leftrightarrow \sum\{P(w, \phi!); (w, \phi!) \in W' \text{ and } M(\phi!), (w, \phi!) \models \phi' \text{ and } M(\phi!), (w, \phi!) \models \psi\} = \sum\{P(w, \phi!); (w, \phi!) \in W' \text{ and } M(\phi!), (w, \phi!) \models \phi'\}$

$\Rightarrow^{*'} \sum\{P(w, \phi!); (w, \phi!) \in W' \text{ and } M(\phi!), (w, \phi!) \models \phi' \to \psi\} = 1$ (1')

$\Leftrightarrow M, w_0 \models [\phi!]C(\phi' \to \psi)$

$\Leftrightarrow \psi \in K^*\phi + \phi'$

**(K\*8)** $\neg\phi' \notin K^*\phi$

$\Leftrightarrow M, w_0 \models \neg[\phi!]C\neg\phi'$

$\Leftrightarrow M(\phi!), (w_0, \phi!) \models \neg C\neg\phi'$

$\Leftrightarrow$ there is a conceived world $(w, \phi!) \in W'$ such that $M(\phi!), (w, \phi!) \models \phi'$ (H')

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

$\sum \{P(w, \phi!); (w, \phi!) \in W'$ and $M(\phi!), (w, \phi!) \models \phi'\}$

$= \sum \{P(w, \phi!); (w, \phi!) \in W'$ and $M(\phi!), (w, \phi!) \models \phi'$ and $(w, \phi!)$ is conceived $\}$ by (H'), because of (H') and if $v$ is conceived and $v'$ surprising then $P_j(v) + P_j(v') = P_j(v)$ (see Sect.1).

$= \sum \{P(w, \phi!); (w, \phi!) \in W'$ and $M(\phi!), (w, \phi!) \models \phi'$ and $M(\phi!), (w, \phi!) \models \psi$ and $(w, \phi!)$ conceived $\}$ by (1')

$= \sum \{P(w, \phi!); (w, \phi!) \in W'$ and $M(\phi!), (w, \phi!) \models \phi'$ and $M(\phi!), (w, \phi!) \models \psi\}$ by (H').

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

∎