What You Should Believe

Guido Boella [a]   Célia da Costa Pereira [b]   Gabriella Pigozzi [c]

Andrea Tettamanzi [b]   Leendert van der Torre [c]

[a] *Università degli Studi di Torino, Dipartimento di Informatica*
guido@di.unito.it
[b] *Università degli Studi di Milano, Dipartimento di Tecnologie dell'Informazione*
pereira@dti.unimi.it, andrea.tettamanzi@unimi.it
[c]*University of Luxembourg, Computer Science and Communication (CSC)*
{gabriella.pigozzi,leon.vandertorre}@uni.lu

Abstract

This paper presents and discusses a normative approach to indeterministic belief revision. An indeterministic belief revision operator assumes that, when an agent is confronted with a new piece of information, it can revise its belief base in more than one way. We propose the epistemic norm that the agent's normative goals should play a role in the choice of one of the available revision alternatives. Properties of the new belief revision mechanism are investigated too.

1   Introduction

Norms and obligations are increasingly being introduced in Multiagent Systems, in particular to meet the coordination needs of open systems where heterogeneous agents interact with each other. Introducing norms raise the issue, however, of the interaction between obligations and other mental attitudes like beliefs, goals, and intentions. While the relation between obligations and motivational attitudes is being studied [2, 7], the relation between beliefs and obligations is still unclear. In this paper we study the role of obligations in the task of revising the agent's beliefs under the light of new information. Revising the beliefs can lead to a situation where a choice among different alternatives cannot be made on the basis of the available information. However, obligations can force an agent to choose among the equally likely alternatives, in order not to miss opportunities.

For example, consider a politician subject to the obligation to reduce deficit, for example due to a decision of the EU or the IMF, who believes that

A) Blocking enrollment leads to a decrease in spending

B) A decrease of investment in infrastructures leads to a decrease in spending

C) A decrease in spending leads to a reduction of deficit

Therefore, his plan to meet his obligation is either to block the enrollment, or to decrease investment in infrastructures.

Now, suppose that someone very trustworthy and well-reputed convinces him that blocking enrollment does not lead to a reduction of deficit. Beliefs A and C cannot hold together anymore, and he has to give up one of them.

If he gives up A, then he still has another possibility to reduce the deficit, because he can decrease spending by decreasing investment in infrastructures. However, if he gives up C, then he does not have any possibility to achieve the reduction of deficit. Indeed,

1. Let us first assume that A is factually wrong, whereas C is true. If he chooses to retain (wrong) belief A and to reject C, then he will do nothing and he will not succeed in reducing deficit. But, had he kept his belief in C and rejected A, then he could have decreased investment in infrastructures in order to decrease spending, and therefore he could have met his obligation to reduce deficit. To conclude, by choosing to maintain A, he risks to miss an opportunity to meet his obligation.

2. Let us now assume that A is actually true and C is wrong. If he chooses to keep (wrong) belief C, then he will decrease spending, but he will not achieve the normative goal of reducing deficit. However, even if he had chosen the right revision, i.e., to retain A and reject C, there was no way for him to achieve his goal of reducing deficit. To conclude, by choosing C (wrong), he believes that he could achieve a goal when he could not, so he will be disappointed for trying in vain, but at least he tried.

The moral of the story is that, if our politician is interested only in meeting the obligation (and there are no other normative goals relevant for him), choosing to maintain C is the only prescribed choice. This is because, independently of C being right or wrong, by choosing that belief he will be at least as well off. Moreover, in one situation — the former — he will be better off if he chooses C than if he chooses A. Summarizing, he should drop A, because that way, he keeps all possibilities to achieve his normative goal open.

We can formalize the above example, by defining the following atomic propositions:

b blocking enrollment; d reduce deficit;
s decrease spending; i decrease investment in infrastructures.

The belief base before being convinced that blocking enrollment does not lead to a reduction of deficit $(\neg(b \supset d))$ would contain the three formulas $b \supset s$, $i \supset s$, and $s \supset d$. The agent has to, first of all, reduce deficit, d, and, if possible, not decrease investment in infrastructures, $\neg i$. Adding $\neg(b \supset d)$ to the beliefs would make them inconsistent. Therefore, he has to revise his beliefs by giving up either $b \supset s$ or $s \supset d$. The choice he makes may depend on the obligations he can meet in the alternatives: if he gives up $b \supset s$, his plan will be to decrease investment in infrastructures, so he will not achieve $\neg i$, but might succeed in achieving d; if he gives up $s \supset d$, his plan will be to do nothing, so he will certainly not achieve d, but he will fulfill $\neg i$. Depending on the punishment he receives for violating his obligation to achieve d or $\neg i$, he could prefer one or the other alternative.

We use the deficit reduction example as a running example throughout the paper.

The choice among belief bases is distinct from other decision problems, due to the possibility of wishful thinking. Consider for example that it is obligatory that blocking the enrollment leads to decrease spending $(b \supset d)$, and that this obligation is more important than the obligation of reducing deficit (d). What should the agent do? At least in a naive approach, he could reason by cases as follows. Assume I choose $b \supset s$: in that case I believe that accomplishing the obligation of blocking enrollment leads to a decrease in spending. Assume I choose $s \supset d$: in that case I believe I am still able to achieve the normative goal of reducing deficit. Since b is more important than d, I choose $b \supset s$.

Instead, the idea of this paper is inspired by the notion of *conventional wisdom* (CW) as introduced by economist John Kenneth Galbraith:

> We associate truth with convenience, with what most closely accords with self-interest and personal well-being. [9]

That is, CW consists of "ideas that are convenient, appealing". This is the rationale for keeping them. One basic brick of CW could then be the fact that some ideas are maintained because they maximize the goals that the agents (believe) they can achieve. This work may be seen as an initial attempt to formally capture the concept of a CW agent. In the following we provide a logical framework that models how a CW agent revises its beliefs under its obligations.

The paper is structured as follows. In Section 2 we introduce the aim of this paper, the used methodology and particular challenges encountered. In Section 3 we introduce the agent theory we use in our approach, and in Section 4 we introduce an indeterministic belief change operator in this agent theory. In Section 5 we define the choice among beliefs as a decision problem in the agent theory. Section 6 concludes.

2   Aim, Methodology and Challenges

The research problem of this paper is to develop a formal model to reason about the kind of choices among belief bases discussed in the previous section, and to generalize the example above in case of additional beliefs, multiple normative goals with different importance, conditional obligations, a way to take violated goals into account, and so on.

We use a combination of the framework of belief revision together with a qualitative decision theory. Classical approaches to belief revision assume that, when an agent revises its belief base in view of new

input, the outcome is well-determined. This picture, however, is not realistic. When an agent revises its beliefs in the light of some new fact, it often has more than one available alternative. Approaches to belief revision that do not stipulate the existence of a single revision option are called *indeterministic* [10]. In this paper we suggest that one possible policy an agent can use in order to choose among available alternatives is to check the effect of the different revisions on the agent's set of goals.

Moreover, for the qualitative decision theory we are inspired by agent theories such as the BOID architecture [2, 7], the framework of goal generation in 3APL as developed by van Riemsdijk and colleagues [12], and [5]. In particular, our agent model is based on one of the versions of 3APL, because the belief base in the mental state of a 3APL agent is a consistent set of propositional sentences, just like in the framework of belief revision. However, we do not pay attention to how goals are generated and how their achievability (plan existence) is established. That is because we do not include "goal-adoption rules" or "practical reasoning rules" representing which action to choose in a particular state. We assume that there is a planning module, which would take a set of goals, actions, and an initial world state representation as input and produce a solution plan as output. This planning module might rely on any propositional AI planner: as in object-oriented programming, we *encapsulate* the planner within a well-defined interface and overlook the implementation details of how a solution plan is found. This is in line with, on one hand, the BOID architecture [2], where the planning component is kept separate from the remainder of agent deliberation, and, on the other hand with the works of Móra and colleagues describing the relationship between propositional planning algorithms and the process of means-end reasoning in BDI agents. In the paper [11], it is shown how the mental state of an agent can be mapped to the STRIPS [8] notation forth and back. This relation has been done on an abstract BDI interpreter named X-BDI [4].

In other words, we model the choice among belief bases essentially as a decision problem, that is, as a choice among a set of alternatives. We do not use classical decision theory (utility function, probability distribution, and the decision rule to maximize expected utility), but a qualitative version based on maximizing achieved goals and minimizing violated goals in an abstract agent theory (see, e.g., [6] for various approaches to formalize the decision process of what an agent should do), because such qualitative decision theories include beliefs and therefore are easier to combine with the theory of belief revision. However, what precisely are the alternatives?

An indeterministic belief revision operator associates multiple revision options to a belief base that turns out to be inconsistent as a consequence of a new piece of information. Our revision mechanism selects the revision alternative that allows the agent to maximize its achievable goals. However, it will not always be possible to select exactly one revision alternative. For example, there may be one most important goal set but two revision alternatives that lead the agent to achieve it. In this case, the two belief revision candidates are said to be equivalent. In Section 5.3 we will provide conditions under which a revision for a CW agent is deterministic, that is, when our revision operator can select exactly one revision alternative.

Besides the issue of wishful thinking, another complicating factor when choosing among belief bases in the context of conditional obligation rules, is that a maximization of goals may lead to a meta-goal to derive obligations by choosing revisions where you believe that the condition is true and the obligation applies. However, deriving goals by itself does not have to be desirable. In contrast, it may even be argued that fewer goals are better than more goals, as you risk to violate goals and become unhappy. We therefore take also goal violations into account.

3   An Abstract Agent Theory

In this section, we represent the formalism which is used throughout the paper.

3.1   A Brief Introduction to AI Planning and Agent Theory

Any agent, be it biological or artificial, must possess knowledge of the environment it operates in, in the form of, e.g., *beliefs*. Furthermore, a necessary condition for an entity to be an *agent* is that it acts. We consider only external motivations like *obligations* and not internal motivations like desires. For artificial agents, obligations may be the purposes an agent was created for.

Obligations are necessary conditions for action, but not sufficient ones. When an obligation is met by other conditions that make it possible for an agent to act, that obligation becomes a normative *goal* [3].

The reasoning side of acting is known as practical reasoning or deliberation, which may include *planning*. Planning is a process that chooses and organizes actions by anticipating their expected effects with the

purpose of achieving as good as possible some pre-stated objectives or goals.

3.2 Beliefs, Obligations, and Goals

The basic components of our language are *beliefs* and *obligations*. Beliefs are represented by means of a *belief base*. A belief base is a finite and consistent set of propositional formulas describing the information the agent has about the world and internal information. Obligations are represented by means of an *obligation base*. An obligation base consists of a set of propositional formulas which represent the situations the agent has to achieve. However, unlike the belief base, an obligation base may be inconsistent, e.g., $\{p, \neg p\}$.

Definition 1 (Belief Base B and Obligation Base O) *Let $\mathcal{L}$ be a propositional language with $\top$ a tautology, and the logical connectives $\land$ and $\neg$, with the usual meaning, and the non material implication $\supset$. The agent's belief base B is a consistent finite set such that $B \subseteq \mathcal{L}$. B can also be represented as the conjunction of its propositional formulas. The agent's obligation base is a possibly inconsistent finite set of sentences denoted by O, with $O \subseteq \mathcal{L}$.*

We define two modal operators Bel and Obl such that, for any formula ϕ of $\mathcal{L}$, $\mathrm{Bel}\phi$ means that ϕ is believed whereas $\mathrm{Obl}\phi$ means that the agent has obligation ϕ. Since in this paper the belief and obligation bases of an agent are completely separated, there is no need to nest the operators Bel and Obl.

Definition 2 (Obligation-Adoption Rule) *An obligation-adoption rule is a triple $\langle \phi, \psi, \tau \rangle \in \mathcal{L} \times \mathcal{L} \times \mathcal{L}$ whose meaning is: if $\mathrm{Bel}\phi$ and $\mathrm{Obl}\psi$, then τ will be adopted as an obligation as well.*

The set of obligation-adoption rules of the agent is denoted by R. If $\langle \phi, \psi, \tau \rangle \in R$ and $\exists \phi', \psi', \tau'$ such that $\phi \leftrightarrow \phi', \psi \leftrightarrow \psi', \tau \leftrightarrow \tau'$, then $\langle \phi', \psi', \tau' \rangle \in R$.

Normative goals, in contrast to obligations, are represented by consistent obligation sets.

Definition 3 (Candidate Goal Set) *A candidate goal set is a consistent subset of O.*

3.3 Mental State Representation

We assume that an agent is equipped with three components:

- belief base $B \subseteq \mathcal{L}$;

- obligation base: $O \subseteq \mathcal{L}$;

- obligation-adoption rule set R.

The mental state $\mathcal{S}$ of an agent is completely described by a triple $\mathcal{S} = \langle B, O, R \rangle$. In addition, we assume that each agent is provided with a problem-dependent function $\mathcal{V}$, a goal selection function G, and a belief revision operator $*$, as discussed below. In our deficit reduction example, we have:

$$
\begin{aligned}
B &= \{b \supset s, i \supset s, s \supset d\}, \\
O &= \{d, \neg i\}, \\
R &= \{\langle \top, \top, d \rangle; \langle \top, \top, \neg i \rangle\}.
\end{aligned}
$$

The semantics we adopt for the belief and obligation operators are standard.

Definition 4 (Semantics of Bel operator) *Let $\phi \in \mathcal{L}$, $\langle B, O, R \rangle \models \mathrm{Bel}\phi \Leftrightarrow B \models \phi$.*

Definition 5 (Semantics of Obl operator) *Let $\phi \in \mathcal{L}$, $\langle B, O, R \rangle \models \mathrm{Obl}\phi \Leftrightarrow \exists$ a maximal (with respect to set inclusion) consistent subset $O' \subseteq O$ such that $O' \models \phi$.*

An agent is obliged to try and manipulate its surrounding environment to fulfill its normative goals. In general, given a problem, not all normative goals are *achievable*, i.e. it is not always possible to construct a plan for each goal. The goals which are not achievable or those which are not chosen to be achieved are called *violated normative goals*. Hence, we assume a problem-dependent function $\mathcal{V}$ that, given a belief base B and a goal set $O' \subseteq O$, returns a set of couples $\langle O^a, O^v \rangle$, where O^a is a maximal subset of achievable normative goals and O^v is the subset of violated normative goals and is such that $O^v = O' \setminus O^a$. Intuitively, by considering violated normative goals we can take into account, when comparing candidate normative goal sets, what we lose from not achieving normative goals.

3.4 Comparing Normative Goals and Sets of Normative Goals

The aim of this section is to illustrate a qualitative method for normative goal comparison in the agent theory. More precisely, we define a qualitative way in which an agent can choose among different sets of candidate normative goals. Indeed, from an obligation base O, several candidate normative goal sets O_i, $1 \leq i \leq n$, may be derived. How can an agent choose among all the possible O_i? It is unrealistic to assume that all normative goals have the same priority. We use the notion of importance of obligations to represent how relevant each normative goal should be for the agent depending, for instance, on the punishment for violating the obligations. The idea is that an agent should choose a set of candidate goals which contains the greatest number of achievable normative goals (or the least number of violated normative goals).

We assume we dispose of a total pre-order $\succeq$ over an agent's obligations. In the example, you have to reduce, in the first place, deficit and, if possible, you should not decrease investments in infrastructures. Therefore, d is more important than $\neg i$, in symbols $d \succeq \neg i$.

The $\succeq$ relation can be extended from goals to sets of goals. We have that a goal set O_1 is more important than another one O_2 if, considering only the goals occurring in either set, the most important goals are in O_1. Note that $\succeq$ is connected and therefore a total pre-order, i.e., we always have $O_1 \succeq O_2$ or $O_2 \succeq O_1$.

Definition 6 (Comparing Sets of Normative Goals)
 The normative goal set O_1 is at least as important as the normative goal set O_2, denoted $O_1 \succeq O_2$ iff the list of obligations in O_1 sorted by decreasing importance is lexicographically greater than the list of obligations in O_2 sorted by decreasing importance. If $O_1 \succeq O_2$ and $O_2 \succeq O_1$, O_1 and O_2 are said to be indifferent, denoted $O_1 \sim O_2$.

In our example, it is easy to verify that $\{d, \neg i\} \succ \{d\} \succ \{\neg i\} \succ \emptyset$. However, we also need to be able to compare the mutual exclusive subsets (achievable and violated goals) of the considered candidate goal, as defined below.

3.5 Comparing Couples of Goal Sets

We propose two methods to compare couples of goal sets.

3.5.1 The *Direct* Comparison $\succeq_D$

Given the $\succeq_D$ criterion, a couple of goal sets $\langle O_1^a, O_1^v \rangle$ is at least as important as the couple $\langle O_2^a, O_2^v \rangle$, noted $\langle O_1^a, O_1^v \rangle \succeq_D \langle O_2^a, O_2^v \rangle$ iff $O_1^a \succeq O_2^a$ and $O_1^v \preceq O_2^v$.
$\succeq_D$ is reflexive and transitive but partial. $\langle O_1^a, O_1^v \rangle$ is strictly more important than $\langle O_2^a, O_2^v \rangle$ in two cases:

1. $O_1^a \succeq O_2^a$ and $O_1^v \prec O_2^v$, or

2. $O_1^a \succ O_2^a$ and $O_1^v \preceq O_2^v$.

They are indifferent when $O_1^a = O_2^a$ and $O_1^v = O_2^v$. In all the other cases, they are not comparable.

3.5.2 The *Lexical* Comparison $\succeq_{Lex}$

Given the $\succeq_{Lex}$ criterion, a couple of goal sets $\langle O_1^a, O_1^v \rangle$ is at least as important as the couple $\langle O_2^a, O_2^v \rangle$ (noted $\langle O_1^a, O_1^v \rangle \succeq_{Lex} \langle O_2^a, O_2^v \rangle$) iff $O_1^a \sim O_2^a$ and $O_1^v \sim O_2^v$; or there exists a $\phi \in \mathcal{L}$ such that both the following conditions hold:

1. $\forall \phi' \succeq \phi$, the two couples are indifferent, i.e., one of the following possibilities holds: (a) $\phi' \in O_1^a \cap O_2^a$; (b) $\phi' \notin O_1^a \cup O_1^v$ and $\phi' \notin O_2^a \cup O_2^v$; (c) $\phi' \in O_1^v \cap O_2^v$.

2. Either $\phi \in O_1^a \setminus O_2^a$ or $\phi \in O_2^v \setminus O_1^v$.

$\succeq_{Lex}$ is reflexive, transitive, and total.

3.6 Defining the Goal Set Selection Function

In general, given a set of obligations O, there may be many possible candidate goal sets. An agent in state $\mathcal{S} = \langle B, O, R \rangle$ must select precisely one of the most important couples of achievable and violated goals.

Let us call G the function which maps a state $\mathcal{S}$ into the couple $\langle O^a, O^v \rangle$ of goal sets selected by an agent in state $\mathcal{S}$. G is such that, $\forall \mathcal{S}$, if $\langle \bar{O}^a, \bar{O}^v \rangle$ is a couple of goal sets, then $G(\mathcal{S}) \succeq \langle \bar{O}^a, \bar{O}^v \rangle$, i.e., a rational agent always selects one of the most preferable couple of candidate goal sets [5].

4 Situating the Problem: Indeterministic Belief Change

"Most models of belief change are deterministic. Clearly, this is not a realistic feature, but it makes the models much simpler and easier to handle, not least from a computational point of view. In indeterministic belief change, the subjection of a specified belief base to a specified input has more than one admissible outcome.

Indeterministic operators can be constructed as sets of deterministic operations. Hence, given n deterministic revision operators $*_1, *_2, \ldots, *_n$, $* = \{*_1, *_2, \ldots, *_n\}$ can be used as an indeterministic operator."

Let us consider a belief base B and a new belief β. The revision of B in light of β is simply:

$$B * \beta \in \{B *_1 \beta, B *_2 \beta, \ldots B *_n \beta\}. \tag{1}$$

More precisely, revising the belief base B with the indeterministic operator $*$ in light of the new belief β leads to one of the n belief revision results:

$$B * \beta \in \{B_\beta^1, B_\beta^2, \ldots B_\beta^n\}, \tag{2}$$

where B_β^i is the i-th possible belief revision result.

Applying the operator $*$ is then equivalent to applying one of the virtual operators $*_i$ contained in its definition. While the rationality of an agent does not suggest any criterion to prefer one revision over the others, a defining feature of a CW agent is that it will choose which revision to adopt based on the consequence of that choice. One important consequence is the set of goals the agent will decide to pursue.

In our deficit reduction example, $\beta = \text{Bel}(\neg(b \supset d))$, and

$$B * \beta \in \left\{ \begin{array}{rcl} B_\beta^1 & = & \{\neg(b \supset d), s \supset d, i \supset s\}, \\ B_\beta^2 & = & \{\neg(b \supset d), b \supset s, i \supset s\} \end{array} \right\}. \tag{3}$$

In the next section we propose some possible ways to tackle the problem of choosing one of the revision options.

5 Belief Revision as a Decision Problem

By considering an indeterministic belief revision, we admit $B * \beta$ to have more than one possible result. In this case, the agent must select (possibly) one among all possible revisions. Among the possible criteria for selection, one is to choose the belief revision operator for which the goal set selection function returns the most important goal set. In other words, selecting the revision amounts to solve an optimization problem.

5.1 Indeterministic State Change

The indeterminism of belief revision influences the obligation-updating process. In fact, the belief revision operator is just a part of the state-change operator, which is indeterministic as well, as a consequence of the indeterminism of belief revision. Therefore, $\mathcal{S}_\beta \in \{\mathcal{S}_\beta^1, \mathcal{S}_\beta^2, \ldots, \mathcal{S}_\beta^n\}$, where $\mathcal{S}_\beta^i = \langle B_\beta^i, O_\beta^i, R \rangle$.

Which goal set is selected by an agent depends on G:

$$G(\mathcal{S}_\beta) \in \{G(\mathcal{S}_\beta^1), G(\mathcal{S}_\beta^2), \ldots, G(\mathcal{S}_\beta^n)\}. \tag{4}$$

In the example, $G(\mathcal{S}_\beta) \in \{G(\mathcal{S}_\beta^1), G(\mathcal{S}_\beta^2)\}$, where $G(\mathcal{S}_\beta^1) = \langle \{d\}, \{\neg i\} \rangle$ and $G(\mathcal{S}_\beta^2) = \langle \{\neg i\}, \{d\} \rangle$. The following table summarizes the possibilities the agent may face when choosing between the two alternative revisions.

reality → ↓ beliefs	$\not\models b \supset s$ $\models s \supset d$	$\models b \supset s$ $\not\models s \supset d$
B_β^1 decrease investment in infrastructures	d is achieved $\neg i$ is not achieved	no obligation is met
B_β^2 do nothing	d is not achieved $\neg i$ is achieved	

A traditional rational agent could not choose one of the $G(\mathcal{S}_\beta^i)$ because they are incomparable. Now, for a CW agent, $G(\mathcal{S}_\beta) \in \mathrm{I}\{G(\mathcal{S}_\beta^1), G(\mathcal{S}_\beta^2), \ldots, G(\mathcal{S}_\beta^n)\}$, where $\mathrm{I}(S)$ denotes the most important set of S defined as follows:

Definition 7 (Important Set I**)** *Given two sets S and X such that $S \subseteq X$, and given an importance relation $\succeq$ over X, the most important set of S is $\mathrm{I}(S) = \{x \in S : \neg \exists x' \in S, x' \succ x\}$.*

5.2 Choosing a Revision

Choosing the most important revision option is not a trivial operation. We can distinguish two situations:

- there is just one most important couple of goal sets $\langle \bar{O}^a, \bar{O}^v \rangle$, but more than one alternative options lead to $\langle \bar{O}^a, \bar{O}^v \rangle$;

- there is no unique most important couple of goal sets; that is, there are different couples of goal sets $\langle O_1^a, O_1^v \rangle, \ldots, \langle O_m^a, O_m^v \rangle$, none of which is strictly more important than the others, i.e., for all $i, j \in \{1, \ldots, m\}$, $\langle O_i^a, O_i^v \rangle \succeq \langle O_j^a, O_j^v \rangle$.

Definition 8 (Equivalent Belief Revision Candidates) *A belief revision candidate B_β^1 is equivalent to another belief revision candidate B_β^2 (denoted by $B_\beta^1 \approx B_\beta^2$), if and only if $G(\mathcal{S}_\beta^1) \succeq G(\mathcal{S}_\beta^2)$ and $G(\mathcal{S}_\beta^2) \succeq G(\mathcal{S}_\beta^1)$.*

It is easy to verify that $\approx$ is a standard equivalence relation, i.e., reflexive, symmetric, and transitive.

The choice of which revision outcome to adopt may thus be deterministic or indeterministic. It is indeterministic in the two cases presented above. More precisely, the choice depends on the importance relations over the couples of goal sets, which determine the equivalence between revision candidates:

- if $\|\mathrm{I}\{G(\mathcal{S}_\beta^1), G(\mathcal{S}_\beta^2), \ldots, G(\mathcal{S}_\beta^n)\}\| = 1$, i.e., the equivalent class of an important belief revision is a singleton and, if there is no i, j such that $G(\mathcal{S}_\beta^i) = G(\mathcal{S}_\beta^j)$, the choice of the belief operator is obviously deterministic;

- if $\|\mathrm{I}\{G(\mathcal{S}_\beta^1), G(\mathcal{S}_\beta^2), \ldots, G(\mathcal{S}_\beta^n)\}\| = 1$, and there is at least a couple i, j such that $G(\mathcal{S}_\beta^i) = G(\mathcal{S}_\beta^j)$, the choice is indeterministic, but also indifferent;

- if $\|\mathrm{I}\{G(\mathcal{S}_\beta^1), G(\mathcal{S}_\beta^2), \ldots, G(\mathcal{S}_\beta^n)\}\| > 1$, the choice is indeterministic.

It is important to notice that an agent that has to choose between $G(\mathcal{S}_\beta^i)$ and $G(\mathcal{S}_\beta^j)$ is in a different situation than an agent that has to randomly choose among a number of competing revisions. But, when an agent must choose between two revision options, it knows that, no matter which revision it chooses, the outcome does not change. In such a context, a random choice becomes a rational option.

Proposition 1 *Let $*$ be an indeterministic belief operator, and n be the number of possible belief revisions candidate. We have $1 \leq \|\mathrm{I}\{G(\mathcal{S}_\beta^1), G(\mathcal{S}_\beta^2), \ldots, G(\mathcal{S}_\beta^n)\}\| \leq n$.*

5.3 Conditions for Determinism of a CW Agent

Traditional indeterministic belief revision approaches allow for the result of belief revision to be indeterminate in the sense that there may be many possible revision alternatives that are equally rational. Our proposal builds on the idea that what an agent wishes to achieve can play a role in the choice of which beliefs to reject and which beliefs to retain. The example we have been using in this paper also tries to capture the intuition that an agent who behaves in this manner is rational. Our richer model can distinguish one revision alternative from the other depending on the effect that each option has on the agent's goal set. Hence, under certain conditions, the choice among several revision alternatives can be reduced to one. This is what we

want to investigate now, that is we want to investigate the conditions under which a revision for a CW agent is deterministic even if an indetermistic revision operator is used, i.e., $\|\mathrm{I}\{G(\mathcal{S}_\beta^i)\}_{i=1,\ldots}\| = 1$ and, for all i, j, $G(\mathcal{S}_\beta^i) \neq G(\mathcal{S}_\beta^j)$.

Observation 1 $B * \beta$ *is deterministic in state* $\mathcal{S} = \langle B, O, R \rangle$, *iff no two alternative revisions are equivalent, i.e., for all* i, j, $B_\beta^i \not\approx B_\beta^j$, *and* $\succeq$ *is total.*

Proposition 2 *A sufficient condition for no two alternative revisions,* B_β^i *and* B_β^j, *being equivalent is that*

1. *for all* i, j, $G(\mathcal{S}_\beta^i) \neq G(\mathcal{S}_\beta^j)$ *and* $G(\mathcal{S}_\beta^i)$ *and* $G(\mathcal{S}_\beta^j)$ *are comparable;*

2. *the importance relation on goals is strict, i.e., for all* $\phi, \phi' \in G(\mathcal{S}_\beta)^a \cup G(\mathcal{S}_\beta)^v$, $\phi \neq \phi'$, $\phi \succeq \phi' \Rightarrow \phi' \not\succeq \phi$.

Proof: From Hypothesis 1 and 2, by applying Definition 6, we obtain $B_\beta^i \not\approx B_\beta^j$. Therefore, no two alternative revisions can be equivalent. $\square$

6 Conclusions

A new framework, inspired by the concept of *conventional wisdom*, aiming at dealing with indeterminism in belief revision has been proposed. While a traditional agent would not be able to choose among multiple revision candidates in indeterministic belief revision, a CW agent evaluates the effects the different revision options have on its goals and selects the revision which maximizes its achievable goals. Fundamental definitions and properties of such belief revision mechanism have been given.

References

[1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *J. Symb. Log.*, 50(2):510–530, 1985.

[2] J. Broersen, M. Dastani, J. Hulstijn, and L. van der Torre. Goal generation in the BOID architecture. *Cognitive Science Quarterly Journal*, 2(3–4):428–447, 2002.

[3] C. Castelfranchi and R. Conte. Commitment: from intentions to groups and organizations. In *Proc. of Autonomous Agents Workshop on Norms and Institutions in Multi-Agent Systems*, Barcelona, 2000.

[4] Michael da Costa Móra, José Gabriel Pereira Lopes, Rosa Maria Vicari, and Helder Coelho. Bdi models and systems: Bridging the gap. In *ATAL*, pages 11–27, 1998.

[5] C. da Costa Pereira and A. Tettamanzi. Towards a framework for goal revision. In Wim Vanhoof Pierre-Yves Schobbens and Gabriel Schwanen, editors, *BNAIC-06, Proceedings of the 18th Belgium-Netherlands Conference on Artificial Intelligence*, pages 99–106. University of Namur, 2006.

[6] M. Dastani, J. Hulstijn, and L. van der Torre. How to decide what to do? *European Journal of Operational Research*, 160(3):762–784, february 2005.

[7] M. Dastani and L. van der Torre. What is a normative goal? towards goal-based normative agent architectures regulated agent-based systems. LNAI 2934, pages 210–227. Springer, 2004.

[8] Richard Fikes and Nils J. Nilsson. Strips: A new approach to the application of theorem proving to problem solving. *Artif. Intell.*, 2(3/4):189–208, 1971.

[9] John K. Galbraith. *The Affluent Society*. Houghton Mifflin, Boston, 1958.

[10] S. Lindström and W. Rabinowicz. Epistemic entrenchment with incomparabilities and relational belief revision. In A. Fuhrmann and M. Morreau, editors, *The Logic of Theory Change*, pages 93–126. 1991.

[11] Felipe Rech Meneguzzi, Avelino Francisco Zorzo, and Michael da Costa Móra. Mapping mental states into propositional planning. In *AAMAS*. ACM Press, 2004.

[12] M. Birna van Riemsdijk. *Cognitive Agent Programming: A Semantic Approach*. PhD thesis, University of Utrecht, 2006.