# Contradicting Beliefs and Communication[*]

Jean-Marc Tallon[†] Jean-Christophe Vergnaud[‡] Shmuel Zamir[§]

January 2003

**Abstract**

We address the issue of the representation as well as the evolution of (possibly) mistaken beliefs. We develop a formal setup (a mutual belief space) in which agents might have a mistaken view of what the model is. We then model a communication process, by which agents communicate their beliefs to one another. We define a revision rule that can be applied even when agents have contradictory beliefs. We study its properties and, in particular, show that, when mistaken, agents do not necessarily eventually agree after communicating their beliefs. We finally address the dynamics of revision and show that when beliefs are mistaken, the order of communication may affect the resulting belief structure.

## 1 Introduction

It is a fact of life that we sometimes do hold mistaken beliefs. We might be wrong about some "objective" fact such as the height of the Mont Blanc (which was recently found to be a few meters higher than previously believed) but we might also be wrong about others' beliefs ("I believed you believed I was already gone") or others' beliefs about our own beliefs ("I believed you believed I believed we would meet at noon rather than at 1 p.m.") and so on. By and large, economic theory has generally ignored this fact, despite the potential interest of allowing for such a possibility in economic modelling. For instance, central bankers' announcements will a priori have very different effects according to whether investors' beliefs about the state of the economy and about what the central bank believes are correct or not. Shareholders might also react differently about disclosure of information by firms, according to whether such information was expected or caught them by surprise. The communication by scientific agencies of discoveries might also be interpreted differently by agents according to how confident they are that their initial beliefs, now contradicted, were right.

[*]We thank Bernard Walliser for useful comments and discussion. We also thank the audience at RUD 2002 in Gif sur Yvette and participants of the workshop on Belief Revision held at LIP6, U. Paris VI. Financial support from the French Ministry of Research (ACI "Cognitique") is gratefully acknowledged.

[†]CNRS–EUREQua, jmtallon@univ-paris1.fr

[‡]CNRS–EUREQua, vergnaud@univ-paris1.fr

[§]CNRS–EUREQua and CREST–LEI, zamir@ensae.fr

In this paper, we address the issue of the representation as well as the evolution of (possibly) mistaken beliefs. The formal setup we develop allows one to model situations in which agents do not have the same view of what is the actual model of the economy. Hence, after communication of each other's beliefs, they might have to deal with surprises or unforeseen contingencies: agent $i$ may be proven wrong in his beliefs about what $j$ thinks the model is after he ($i$) hears $j$'s announcement. As a consequence, a lot of the intuition one has formed in the standard case (i.e., with no mistake) does not hold in this more general setting. For instance, communication does not necessarily lead to agreement and might well lead to a situation in which agents disagree and agree to disagree (i.e., disagreement is "common knowledge"). Thus results such as agreeing to disagree à la Aumann (1976) and the convergence of beliefs as in Geanakoplos and Polemarchakis (1982) have to be reconsidered in situations entailing mistakes (on the underlying state of nature or on some higher order beliefs.)

The main contribution of the paper is to define a revision rule that applies when agents learn of something that they did not believe possible originally. This rule encompasses both cases in which agents do not face contradictions and can therefore simply refine their original beliefs as well as when they have to modify their beliefs so as to acknowledge some fact they thought impossible. The properties of the rule however do differ in these two cases: as mentioned above, disagreement might be an outcome of communication in the presence of contradictions. Another important feature of the revision rule in the general case is that the final situation reached might depend upon the order of announcements. Indeed, when extended to a dynamic setting, the sequence of revisions may not be commutative.

To carry this analysis, we introduce the concept of a mutual belief system, in which beliefs are expressed not in terms of probability but in terms of possibility correspondence: an agent's beliefs at a state $\omega$ are represented by the set of states of the world he considers possible when the state is $\omega$. This allows for a clear-cut definition of what it means to be mistaken: an agent has mistaken beliefs in a given state of the world if he does not consider that state as possible (i.e., when the true world is $\omega$, the agent does not believe $\omega$ is possible.) Mutual belief systems is an instance of what has been introduced in the modal logic literature under the name of Kripke structure. It consists of a complete description of the agents' beliefs about the state of nature, the beliefs of other agents about the state of nature, the beliefs of agents on others' beliefs about the state of nature and so forth. This infinite hierarchy of beliefs can be embedded in a mutual belief system, which has a self reference structure. Mathematically, mutual belief system is the analogue of the concept of belief subspace (with specification of the true state of the world) of Mertens and Zamir (1985) when beliefs are expressed not in terms of probability but in terms of possibility correspondence. We furthermore adopt an interim viewpoint, that is we consider the mutual beliefs of the agents at the *true state of the world*, which is assumed to be given and fixed. Relaxing the assumption (usually referred to as the *truth axiom*) that agents have always correct beliefs brings up interesting issues concerning the definition and characterization of common belief in mutual belief systems, given that different agents might have different models of the world. In particular, such a belief structure with mistakes is not necessarily commonly believed or known by all agents (as it is assumed for instance when a game is given). Our model is therefore that viewed by an outside observer (the analyst).

Having set a general framework, we allow for communication of beliefs among agents (in a non strategic way) and study the revision of beliefs after communication has taken place. When agents' beliefs are correct, it is easy to come up with a revision rule that expresses the fact that agents sharpen their beliefs upon hearing the others'. Essentially, one only need to take the intersection of one's initial beliefs with the announcements of the other. However, defining a sensible revision rule when some agents have mistaken beliefs is less straightforward. Indeed, contradictions between agents' initial beliefs and the announcements of the others have to be dealt with. The crucial question is then, how does an agent whose beliefs are proven wrong by the announcements of some other agent revise her beliefs? Consider the following simple situation: agent 1 believes that the state of nature is $\alpha$ or $\beta$, believes that agent 2 believes that the state of nature is $\alpha$ or $\beta$ and actually believes that this is *common beliefs*. However, agent 2 believes that the state of nature is $\beta$, that agent 1 believes it is $\beta$ and that this is common belief: how should agent 1 (resp. 2) change her beliefs, which are proven wrong when 2 (resp. 1) announces that he believes that $\beta$ is common belief (resp. $\alpha$ or $\beta$ is common belief)? In this simple example, if both agents announce their beliefs, a possible revision would simply be that both agents now believe that the state is $\beta$ and that this is commonly believed.[1] Alternatively, both agents could well hold on to their first order beliefs (1 continues to believe the state of nature is $\alpha$ or $\beta$ and 2 still believes it is $\beta$) and change their higher order beliefs to take into account the announced beliefs: the situation reached would then be that 1 believes $\alpha$ or $\beta$, 2 believes $\beta$ and this situation is common belief (that is, the two agents "agree to disagree".) This latter rule is based on the idea that agents' revised beliefs are "entrenched" in their initial beliefs, an idea that was already put forth in the axiomatic literature on (single agent) belief revision developed by Alchourrón, Gärdenfors and Makinson (1985) (see also Makinson (1985)).

We develop a general revision rule when agents announce truthfully their exact beliefs, that always yields well defined mutual belief systems. Thus, the presence of mistakes is not synonymous of "anything can happen" and in particular the consistency properties embedded in the definition of a mutual belief system (akin to positive and negative introspection) continue to hold. The revision rule is based on the intuition underlying the second revision rule described in the example above that agents will change their beliefs in a "minimal way" when they face a contradiction. This amounts to assume some form of entrenchment of the revised beliefs in the original ones. In particular, in our model, an agent will never believe that a state of nature he initially believed impossible is now possible simply because some other agent announced that

[1] This example will be treated formally as Example 2 in the text.

she believed it possible.

We then proceed to the analysis of the properties of this rule in a static setting. For instance, we give sufficient conditions that ensure that the mutual belief system reached after communication does not entail any disagreement. These conditions are stronger than simply the absence of mistakes in the true state of the world. Essentially, it is necessary that the absence of mistakes be common belief to yield agreement among agents. We next turn to the dynamics of belief revision, when announcements are made sequentially, and uncover a few interesting issues: first, we establish that the mutual belief system eventually reached is one in which first order beliefs are common beliefs. Second, we show through examples that the order of announcements might matter as to where the system converges to. This is an instance in which the intuition one might have when beliefs are not mistaken (in which case the order of communication does not matter) does not generalize.

We believe that the setting developed in this paper will be useful for dealing with issues such as trade based on differences in beliefs, since our setting allows for differences in beliefs that are of a different kind than for instance the ones studied in Morris (1994). In Morris (1994), although agents do not have a common prior, they can update their beliefs according to Bayes' rule and do not have to deal explicitly with contradictions. In our setting, there is not a single way to revise beliefs when faced with a contradiction. We are not the first ones to introduce Kripke structures in economics (see for instance Bacharach (1985), Samet (1990)) nor to relax the truth axiom in such a setting (see Bonanno and Nehring (1998)). However, to the best of our knowledge, the analysis of revision of beliefs in such a setting has not been pursued.

The paper is organized as follows. Section 2 contains the definition and some properties of mutual beliefs systems. Section 3 develops the notion of common beliefs in mutual belief systems. In Section 4, the heart of the paper, we define a belief revision rule and study its properties. Appendix A and B contain some technical material that can be skipped in a first reading. All proofs are gathered in Appendix C.

## 2 Mutual Belief Systems: definition and preliminaries

Let $I = \{1, ..., i, ..., n\}$ be a finite set of agents and $S$ a set of states of nature. A mutual belief system is a representation of agents' beliefs about the state of nature $s$ and about the beliefs of the other agents. Because of this latter aspect, the structure introduced has to be self-referential, as one can see in the following definition.

**Definition 1** *A Mutual Belief System (MBS) is a collection* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *, where* $\Omega$ *is a set, and the following conditions are satisfied:*

*(i)* $s$ *is a mapping from* $\Omega$ *to* $S$*,*

*(ii)* $\forall i \in I$*,* $t_i$ *is a mapping from* $\Omega$ *to* $2^{\Omega}$*,*

*(iii)* $\forall i \in I$*,* $\forall \omega \in \Omega$*,* $\omega' \in t_i(\omega) \Rightarrow t_i(\omega') = t_i(\omega)$*,*

*(iv)* $\omega_0 \in \Omega$*,*

*(v) There does not exist $\Omega' \subsetneq \Omega$ such that $(\Omega', \omega_0, s|_{\Omega'}, (t_i|_{\Omega'})_{i\in I})$ satisfies conditions (i) to (iv).*[2]

An element $(\omega; s(\omega); t_1(\omega), ..., t_n(\omega))$ is called a *state of the world*. Its interpretation is as follows: $\omega$ is the name of the state, $s(\omega)$ is the state of nature in the world $\omega$, $t_i(\omega)$ is the set of states of the world that $i$ considers possible in state $\omega$ (and can also be thought of as "the type" of agent $i$ in state $\omega$). Finally, $\omega_0$ is the true state of the world. Abusing notation slightly we will denote a state of the world $\omega = (s(\omega), t_1(\omega), ..., t_n(\omega))$ since $\omega$ uniquely determines the state of the world.

It is important to note that the definition does *not* require that agents consider $\omega_0$ possible, i.e., $\omega_0$ need not be in $t_i(\omega_0)$. A consequence of allowing mistaken beliefs is that the MBS is not necessarily known by the agents. Embedded in the definition are several assumptions about the nature of the situations we model. First, we assume a form of consistency of the beliefs: $(iii)$ of the definition implies that beliefs are partitional (i.e., $\{t_i(\omega)\}_{\omega\in\Omega}$ is a partition of $\Omega_i =: \cup_{\omega\in\Omega} t_i(\omega)$). Note however that $\Omega_i$ is not necessarily equal to $\Omega$. Second, the true state $\omega_0$ is given. Thus, we place ourselves in a situation often referred to as the "interim stage", at which the state of nature and the beliefs of the agents are realized. Third, we assume that the mutual belief system is minimal in the sense that it does not contain a smaller MBS (condition $(v)$). This last condition is equivalent to assuming that the mutual belief system does not contain states that are not deemed possible via a finite sequence of steps of the form "I think that you think that she thinks..." (condition $(v')$ in Proposition 1 below, which will be used repeatedly in the proofs of this paper.) These are states that are "not in the mind" of any player. This does not imply that $\Omega$ is finite. In particular, we could represent in our setup Rubinstein's electronic mail game (Rubinstein (1989)) which necessitates an infinite state space. The restriction imposed is, rather, that among any two given worlds $\omega$ and $\omega'$, the "distance" is finite, i.e., there is finite path that links the two worlds via chains of the form in state $\omega$, $i_1$ believes that $i_2$ believes that ... $i_k$ believes $\omega'$.

**Proposition 1** *Let $(\Omega, \omega_0, s, (t_i)_{i\in I})$ be a collection which satisfies conditions (i) to (iv) of Definition 1. Then condition (v) is equivalent to*

*(v') $\forall\omega \in \Omega \setminus \{\omega_0\}$, there exists a finite sequence, $\{i_k\}_{k=1}^{k=r}$ with $i_k \in I$ for all $k$ such that $\omega \in t_{i_1}(t_{i_2}(...(t_{i_r}(\omega_0))))$ where for any $A \subset \Omega$, $t_i(A) = \cup_{\omega\in A} t_i(\omega)$.*

We now illustrate the concept of an MBS on the well-known example of the three hats.

**Example 1** *Three girls wear hats that can be either red (R) or black (B). Each girl sees the other two girls' hats but does not see the color of her own hat. Assume that the three hats are actually red. Denoting the states of nature by $C_1C_2C_3$ where*

[2] $t_i|_{\Omega'}$ is the restriction of $t_i$ to $\Omega'$, i.e., $t_i|_{\Omega'} : \Omega' \to 2^\Omega$ and $t_i|_{\Omega'}(\omega) = t_i(\omega)$ for all $\omega \in \Omega'$.

*$C_i \in \{R, B\}$ is the color of agent i's hat, we represent this situation by an MBS given by $\Omega = \{\omega_0, ..., \omega_7\}$ where,*

$\omega_0 = (RRR, \{\omega_0, \omega_4\}, \{\omega_0, \omega_2\}, \{\omega_0, \omega_1\})$
$\omega_1 = (RRB, \{\omega_1, \omega_5\}, \{\omega_1, \omega_3\}, \{\omega_0, \omega_1\})$
$\omega_2 = (RBR, \{\omega_2, \omega_6\}, \{\omega_0, \omega_2\}, \{\omega_2, \omega_3\})$
$\omega_3 = (RBB, \{\omega_3, \omega_7\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\})$
$\omega_4 = (BRR, \{\omega_0, \omega_4\}, \{\omega_4, \omega_6\}, \{\omega_4, \omega_5\})$
$\omega_5 = (BRB, \{\omega_1, \omega_5\}, \{\omega_5, \omega_7\}, \{\omega_4, \omega_5\})$
$\omega_6 = (BBR, \{\omega_2, \omega_6\}, \{\omega_4, \omega_6\}, \{\omega_6, \omega_7\})$
$\omega_7 = (BBB, \{\omega_3, \omega_7\}, \{\omega_5, \omega_7\}, \{\omega_6, \omega_7\})$

*The same example, in which it is common knowledge that there is at least one red hat among the three girls is described by $\Omega = \{\omega_0, ..., \omega_6\}$ where,*

$\omega_0 = (RRR, \{\omega_0, \omega_4\}, \{\omega_0, \omega_2\}, \{\omega_0, \omega_1\})$
$\omega_1 = (RRB, \{\omega_1, \omega_5\}, \{\omega_1, \omega_3\}, \{\omega_0, \omega_1\})$
$\omega_2 = (RBR, \{\omega_2, \omega_6\}, \{\omega_0, \omega_2\}, \{\omega_2, \omega_3\})$
$\omega_3 = (RBB, \{\omega_3\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\})$
$\omega_4 = (BRR, \{\omega_0, \omega_4\}, \{\omega_4, \omega_6\}, \{\omega_4, \omega_5\})$
$\omega_5 = (BRB, \{\omega_1, \omega_5\}, \{\omega_5\}, \{\omega_4, \omega_5\})$
$\omega_6 = (BBR, \{\omega_2, \omega_6\}, \{\omega_4, \omega_6\}, \{\omega_6\})$

Although for the sake of simplicity, the examples we give in the paper are abstract examples (i.e., we do not provide an economic interpretation), economic example can be constructed in a similar fashion. Take for instance the situation in which the central bank (agent 1) is correct about the true state of the economy (say, Boom) while investors (agent 2) consider both states (Boom or Recession) possible but believe that the central bank knows the true state of the economy and believes this is *common belief* (a notion we will define precisely in Section 2). Such a situation would be represented (assuming the true state of nature is indeed "Boom") by the following MBS (the central bank is the first agent, investors the second): $\Omega = \{\omega_0, \omega_1\}$ where

$\omega_0 = (Boom, \{\omega_0\}, \{\omega_0, \omega_1\})$
$\omega_1 = (Recession, \{\omega_1\}, \{\omega_0, \omega_1\})$

Expanding a bit on this example, one could also consider a situation in which investors wrongly believe that the central bank is informed of the true state of the economy and the central bank believes that investors believe it is informed of the true state of the economy. Such a situation would be captured by $\Omega = \{\omega_0, \omega_1, \omega_2, \omega_3\}$ where,

$\omega_0 = (Boom, \{\omega_0, \omega_1\}, \{\omega_2, \omega_3\})$
$\omega_1 = (Recession, \{\omega_0, \omega_1\}, \{\omega_2, \omega_3\})$
$\omega_2 = (Boom, \{\omega_2\}, \{\omega_2, \omega_3\})$
$\omega_3 = (Recession, \{\omega_3\}, \{\omega_2, \omega_3\})$

This is an instance of an MBS entailing mistaken beliefs. The next example illustrates a similar instance of mistaken beliefs, which was verbally presented in the introduction.

**Example 2** *Let* $S = \{\alpha, \beta\}$ , $I = \{1, 2\}$ *and* $\Omega = \{\omega_0, \omega_1, \omega_2, \omega_3\}$ *such that:*
$\omega_0 = (\alpha, \{\omega_1, \omega_2\}, \{\omega_3\})$
$\omega_1 = (\alpha, \{\omega_1, \omega_2\}, \{\omega_1, \omega_2\})$
$\omega_2 = (\beta, \{\omega_1, \omega_2\}, \{\omega_1, \omega_2\})$
$\omega_3 = (\beta, \{\omega_3\}, \{\omega_3\})$

This represents a situation in which the true state of nature is $\alpha$, agent 1 believes that it is $\alpha$ or $\beta$ and agent 2 believes that it is $\beta$. Furthermore, 1 believes that 2 believes that the state of nature is $\alpha$ or $\beta$ while 2 believes that 1 believes that the state of nature is $\beta$. In a nutshell, 1 believes that it is common belief that the state of nature is $\alpha$ or $\beta$, while 2 believes that it is common belief that the state is $\beta$.

The restriction imposed by $(v)$ in Definition 1 can be illustrated in Example 2 above: if one were to add a state like $\omega_4 = (\alpha, \omega_0, \omega_4)$ to the MBS in Example 2 (leaving $\omega_0$, $\omega_1$, $\omega_2$, and $\omega_3$ unchanged), then, formally, the system thus obtained is not an MBS since it contains a subset (namely $\{\omega_0, \omega_1, \omega_2, \omega_3\}$) which satisfies conditions $(i)$ to $(iv)$ of Definition 1. Alternatively, one can observe that $\omega_4$ cannot be reached by a finite chain of beliefs as required by condition $(v')$ of Proposition 1. Finally, the definition of an MBS should make it clear that the same epistemic state of the agents could be represented in various ways.

**Example 3** *Let* $S = \{\alpha, \beta\}$ , $I = \{1, 2\}$ *and* $\Omega = \{\omega_0, \omega_1, \omega_2, \omega_3, \omega_4\}$ *such that:*
$\omega_0 = (\alpha, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\})$
$\omega_1 = (\alpha, \{\omega_1, \omega_2\}, \{\omega_1, \omega_2\})$
$\omega_2 = (\beta, \{\omega_1, \omega_2\}, \{\omega_1, \omega_2\})$
$\omega_3 = (\beta, \{\omega_3\}, \{\omega_3, \omega_4\})$
$\omega_4 = (\beta, \{\omega_4\}, \{\omega_3, \omega_4\})$
*An examination of the beliefs represented here reveals that the epistemic situation is the same as the one in Example 2. Hence, we could say that the MBS in Example 2 is a* representation *of the MBS defined above and that the two MBS, capturing the same epistemic situations are* equivalent.

The previous example reveals that a given epistemic situation could be captured by MBS that are formally different. This fact is not bothersome if agents do not make any mistake. However, as we want to study revision in beliefs when agents potentially have initial mistaken beliefs, we have to make sure that "irrelevant" mistakes can be dropped at the outset so as to focus on beliefs that are mistaken in a meaningful way. A simple intuition of why some mistakes are not meaningful is the following: imagine that $\omega_0 \notin t_i(\omega_0)$. This can reflect two very different situations: either the agent is correct in the sense that in $\omega_0$ he believes possible a state $\omega'$ which represents the same beliefs as $\omega$; or the agent is making a mistake in the sense that he is not considering as possible the true state of the world $\omega_0$ (or any state of the world that represents the same epistemic state.)

A way of getting around this difficulty is to define notions of *representation* and *equivalence* of MBS as well as a notion of *minimality* for MBS. This is done in appendix A to which we refer the interested reader. Minimality consists in essentially getting rid of potential redundancies in an MBS. Our definition, whose details can be skipped in a first reading, thus identifies redundancies in Example 3 and suggest to "merge" states $\omega_1$ and $\omega_3$. A minimal MBS is one in which all the redundancies have been removed. In the rest of the paper, we exclusively deal with minimal MBS.

# 3 Common Belief in Mutual Beliefs Systems

In this section, we explore ways of expressing belief properties in MBS. We first define the notion of belief horizon for an agent. It is the set of states of the world that are believed possible by the agent (possibly via links of the form "I believe that you believe this state is possible", or "I believe that you believe that she believes this state is possible",...). We then define common belief and provide a characterization in terms of the agents' belief horizons. We conclude this section by consideration on the notion of correct beliefs.

## 3.1 Belief Horizon and Common Belief

When agents hold mistaken beliefs, they do not necessarily all have the same view of what the model actually is. We introduce here the notion of belief horizon of an agent which is the model the agent has in mind.

**Definition 2** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be an MBS. The* belief horizon *of agent* $i \in I$*, denoted by* $BH_i(\omega_0, t)$*, is the minimal subset* $Y$ *of* $\Omega$ *satisfying:*

(i) $t_i(\omega_0) \subset Y$*,*

(ii) $\forall \omega \in Y$*,* $\forall j \in I$*,* $t_j(\omega) \subset Y$*.*

Thus, $BH_i(\omega_0, t)$ is the smallest "public event" for $i$, i.e., the smallest set such that $i$ believes it and believes that all other agents believe it, believe that others believe that others believe it and so forth. In Example 2, one has $BH_1(\omega_0, t) = \{\omega_1, \omega_2\}$ and $BH_2(\omega_0, t) = \{\omega_3\}$.

**Proposition 2** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be an MBS. For all* $i \in I$*,* $\forall \omega \in \Omega$*,*

$$\omega \in BH_i(\omega_0, t) \Leftrightarrow \exists r \in \mathbb{N}, \exists \{i_k\}_{k=1}^{k=r}, i_k \in I, i_r = i \ s.th. \ \omega \in t_{i_1}(t_{i_2}(\ldots(t_{i_r}(\omega_0)))$$

This Proposition enables us to state a useful property of MBS, namely that an MBS is the union of agents' belief horizons and of the true state (which might not be in any agent's belief horizon).

**Corollary 1** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be an MBS. Then,* $\Omega = \{\omega_0\} \cup (\cup_{i\in I} BH_i(\omega_0, t))$.

The definition of common belief of an event is the usual definition of common knowledge, adapted to our setting: an event is common belief if all agents believe it, all agents believe that all agents believe it and so forth.

**Definition 3** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be an MBS. An event* $E \subset \Omega$ *is* common belief *(CB) if for any* $r \in \mathbb{N}$ *and any sequence* $\{i_k\}_{k=1}^{k=r}$, $i_k \in I$, $t_{i_1}(t_{i_2}(\dots(t_{i_r}(\omega_0))) \subset E$

Note that as an MBS describes a mutual belief structure at a specific, 'true', state of the world, common belief is also defined at that state $\omega_0$. The following proposition can be deduced immediately from Proposition 2.

**Proposition 3** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be an MBS. An event* $E \subset \Omega$ *is common belief if and only if* $BH_i(\omega_0, t) \subset E$ *for all* $i \in I$

This notion of common belief is meaningful for the analyst since, according to $i$'s beliefs, *any event* containing $BH_i(\omega_0, t)$ is CB. As we shall see later, only at the absence of mistakes, CB events have stronger meaning. Recall that in the common knowledge literature (e.g., Geanakoplos (1994)), an event is common knowledge at some state if there exists a public event (or a truism) at that state that is included in the event. The characterization of common belief we provided is similar, although the "public event" might be agent dependent since agents do not necessarily have the same belief horizon.

**Corollary 2** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be an MBS. An event* $E \subset \Omega$ *is common belief if and only if*

$$\cup_{i\in I} BH_i(\omega_0, t) \subset E \subset \Omega = \{\omega_0\} \cup (\cup_{i\in I} BH_i(\omega_0, t))$$

This corollary establishes that in an MBS, at most two events can be common belief. $\Omega$ is always commonly believed (by construction of an MBS), while $\Omega \setminus \{\omega_0\}$ is common belief only if the true state $\omega_0$ does not belong to the belief horizon of any agent. In other words, $\Omega$ is the only common belief event at $\omega_0$ if and only if $\omega_0$ is in the belief horizon of at least one agent.

## 3.2 Correct Mutual Belief Systems

A mutual belief system is correct if agents make no mistake, in the sense they all believe that the true state $\omega_0$ is possible. Correctness, in our framework, is in some sense the analogue of the truth axiom in knowledge systems, which asserts that if an agent knows something then it must be true. However, it is possible to construct examples in which all agents are correct but this is not commonly believed. This stronger notion is captured by the notion of totally correct MBS.

**Definition 4** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be a minimal MBS. An agent* $i \in I$ *has* correct beliefs *if* $\omega_0 \in t_i(\omega_0)$. *The MBS is* correct *if all agents have correct beliefs. The MBS is* totally correct *if* $\omega \in t_i(\omega)$ *for all* $\omega \in \Omega$ *and all* $i \in I$.[3]

In Example 1, the MBS are totally correct, while the MBS of Example 2 is not correct. As mentioned, an MBS can be correct but not totally correct, as illustrated in the following example.

**Example 4** *Let* $S = \{\alpha, \beta\}$ *and* $I = \{1, 2\}$. *Consider* $\Omega = \{\omega_0, \omega_1\}$ *where*
$\omega_0 = (\alpha, \{\omega_0\}, \{\omega_0, \omega_1\})$
$\omega_1 = (\beta, \{\omega_0\}, \{\omega_0, \omega_1\})$

In this Example, the two agents satisfy the truth axiom in the true world $\omega_0$ but agent 2 does not believe that agent 1 satisfies it: 2 believes that in the possible world $\omega_1$, agent 1 is mistaken. Thus there is a difference between situations where all agents satisfy the truth axiom but this fact is not commonly believed (i.e the MBS is correct but not totally correct) and situations which are captured through totally correct MBS where all agents satisfy the truth axiom and this fact is common belief.

If an agent is correct, it is easy to see that his belief horizon contains the belief horizons of all other agents, and his belief horizon is the entire space $\Omega$. A direct corollary of this fact together with Corollary 2 is that if at least one agent is correct, the only common belief event is $\Omega$ itself. Further, since MBS that are correct have the feature that different agents' belief horizons coincide, this common belief horizon is common belief and therefore correctness embeds a kind of agreement among agents about what the model is. Note finally that correctness is sufficient but not necessary for the coincidence of belief horizons of the different agents. Take for instance the following MBS: $\omega_0 = (\alpha, \{\omega_0\}, \{\omega_1\})$ and $\omega_1 = (\beta, \{\omega_0\}, \{\omega_1\})$. There, $BH_1(\omega_0; t) = BH_2(\omega_0; t) = \{\omega_0, \omega_1\}$ while agent 2 is not correct.

# 4 Communication and Revision in Mutual Belief Systems

We are interested in studying the evolution of beliefs when agents can communicate their beliefs to each other and update accordingly. In this Section we provide rules according to which agents revise their beliefs in a communication process. At this stage of our work, we do not allow agents to announce false (or partly false) or even imprecise beliefs. Thus, the analysis will concentrate on the case in which agents announce *truthfully and precisely* their beliefs.

**Definition 5** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be an MBS. A* communication *is a collection* $(t_i(\omega_0))_{i\in I^c}$ *where* $I^c \subset I$.[4]

[3] If the MBS considered were not minimal, the definition should be slightly more general: an MBS is correct if $\forall i \in I$, there exists $\omega \in t_i(\omega_0)$, such that $\omega$ and $\omega_0$ are identical. When the MBS is minimal, this definition and definition 4 coincide.

[4] Thus, we restrict attention to communication that are *full truthful* in the sense that agents who communicate tell the truth, the whole truth and nothing but the truth.

A communication can therefore be identified by $I^c \subset I$, the group of agents who announce their true beliefs (the carrier of the communication). We'll refer to it as *full communication* when $I^c = I$. The restriction that agents announce precisely their true beliefs can be understood as an assumption that the information revealed can be somehow certified. Lying is hence prohibited. The strategic aspects of communication, including the possibility of non truthful announcements are not addressed at this stage of the work. Finally, we will assume in the sequel that it is "common knowledge" that agents announce precisely their true beliefs. This is analogue to the models of Geanakoplos and Polemarchakis (1982) and Bacharach (1985).

We now move on to some attempts to define a revision rule. We show through examples that the most intuitive rules are not adapted to our setting where agents might be mistaken. We then discuss how agents "should" cope with announcements contradicting their initial beliefs, and introduce the notion of an order on the states of the world, reflecting an agent's view of which worlds are "closest" to the true world, after hearing the announcements of the other agents. This enables us to propose a revision procedure and to study its properties. We then generalize it to dynamic communication processes.

## 4.1 Defining a revision procedure: a first attempt

To develop and motivate our definition of a general revision rule, we first look at two simple and intuitive rules. We examine their deficiencies and the domain in which they are adequate. We then generalize them so as to cope with these deficiencies while maintaining their adequate performance.

A first intuition that one might have is simply to assume that each agent takes the announcement of the other agents at face value and hence revises his beliefs by taking the intersection of his initial beliefs with the announcement of the other. A second intuition is that an agent is not directly interested by the content of the announcements but rather by the worlds which are compatible with the announcements, i.e., he considers *the states of the world in which these announcements could have been made*; any other state of the world is eliminated by the revision. To illustrate these two logics, consider the following example:

**Example 5** *Let* $S = \{\alpha, \beta\}$ *and* $I = \{1, 2\}$. *Consider* $\Omega = \{\omega_0, \omega_1, \omega_2\}$ *where*

$\omega_0 = (\alpha, \{\omega_0, \omega_1\}, \{\omega_0, \omega_2\})$

$\omega_1 = (\alpha, \{\omega_0, \omega_1\}, \{\omega_1\})$

$\omega_2 = (\beta, \{\omega_2\}, \{\omega_0, \omega_2\})$

*When* $I^c = I$, *the two revision rules suggested above both yield* $\omega_0 = (\alpha, \{\omega_0\}, \{\omega_0\})$ *that is, both agents learn from the other's announcement that the true state is* $\omega_0$. *However, the process through which one arrives at this MBS is different in the two rules: according to the first intuition, agent* 1 *drops state* $\omega_1$ *because agent* 2 *announced that he does not believe in this state while according to the second intuition, agent* 1 *drops state* $\omega_1$ *because in that state, agent* 2 *would have announced* $\{\omega_1\}$.

This example is actually representative of the class of totally correct MBS in which the two revision rules suggested yield the same, well-defined MBS. Before pointing differences between these two rules, we first introduce them formally.

**Definition 6** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be a totally correct MBS. Given a communication* $I^c$*, the revision of beliefs is the MBS,* $(\Omega^c, \omega_0, s, (t_i^c)_{i\in I})$ *defined by:*[5]

- First revision rule.
  - $\forall i \in I$, $\forall \omega \in \Omega$, $t_i^c(\omega) = t_i(\omega) \cap (\cap_{j\in I^c} t_j(\omega_0))$,
  - $\Omega^c = \{\omega_0\} \cup (\cup_{i\in I} BH_i(\omega_0, t^c))$
- Second revision rule
  - $\forall i \in I$, $\forall \omega \in \Omega$, $t_i^c(\omega) = t_i(\omega) \cap \{\omega' \in \Omega | t_j(\omega') = t_j(\omega_0);\ \forall j \in I^c\}$
  - $\Omega^c = \{\omega_0\} \cup (\cup_{i\in I} BH_i(\omega_0, t^c))$

**Remark 1** *There is a slight abuse of notation in the previous definition, as* $\Omega^c$ *is defined via belief horizons that are, strictly speaking, only defined once* $\Omega^c$ *is given. Furthermore, one can define* $t_i^c$ *only after having defined* $\Omega^c$*. Rigourously, one needs to define* $(\Omega^c, \omega_0, s, (t_i^c)_{i\in I})$ *as follows:*

- $\forall i \in I$, $\forall \omega \in \Omega$, $t_i^{'}(\omega) = t_i(\omega) \cap (\cap_{j\in I^c} t_j(\omega_0))$,
- $\Omega^c = \{\omega_0\} \cup \{\omega \in \Omega | \exists r \in \mathbb{N} \text{ and } \{i_k\}_{k=1}^{k=r}, i_k \in I, \text{ s.th. } \omega \in t_{i_1}^{'}(t_{i_2}^{'}(\ldots(t_{i_r}^{'}(\omega_0)))\}$
- $\forall i \in I$, $t_i^c = t_i^{'}|_{\Omega^c}$.

It is readily verified that the total correctness of the MBS guarantees that the revised beliefs are well defined and hence we have:

**Proposition 4** *The revision of beliefs according to the first and to the second revision rule yield, in a totally correct MBS, the same totally correct MBS.*

When we consider non totally correct MBS, we encounter two kinds of problems. First, the two rules might lead to different MBS.[6] Second, they may not be applicable. Let us examine the first problem on Example 4 where the MBS is correct.

**Example 6** *(Example 4 continued)*

*Let* $I^c = \{1\}$*. Then the first rule leads to* $\omega_0 = (\alpha, \{\omega_0\}, \{\omega_0\})$*, while the second rule leaves the initial MBS unchanged.*

[5] For convenience (or abuse...) of notation, the names of the states of the world in $\Omega^c$ are the same as in $\Omega$ (but with different beliefs of course.)

[6] Although we defined the two rules only for totally correct MBS, it is clear that they are applicable to a wider set of MBS.

We feel that the first rule is not very satisfactory in this example. Indeed, it is as if the first agent managed to convince agent 2 to drop state $\omega_1$: 1 says "I believe the state of nature is $\alpha$" and agent 2 is convinced that he should not consider any more that the state of nature could be $\beta$. Yet, before any communication took place, agent 2 thought that agent 1 could well be mistaken on the state of nature and agent 1's announcement was completely foreseeable by agent 2. Thus, following the first rule leads to admit that agent 2 is influenced by others' announcement even though it was expected and hence is not confident in his own beliefs. The second revision rule leaves the MBS unchanged which looks more reasonable.

In view of this example, we chose to generalize the second intuition rather than the first. This amounts to implicitly assume that one will never abandon one's initial beliefs when they are not proven false. Even if an agent's beliefs are contradicted by the beliefs of another agent, the first agent will not adopt the second agent's beliefs but simply incorporate in his own beliefs the fact that they disagree. There is a sense in which revised beliefs are entrenched in the initial beliefs. This represents situations in which each agent believes that his own expertise is at least as good as the others'.

The second problem, which is faced by the two rules, is that they might be ill-defined for non correct MBS, as shown in the following example.

**Example 7** *Let* $S = \{\alpha, \beta, \gamma\}$ *and* $I = \{1, 2\}$. *Consider* $\Omega = \{\omega_0, \omega_1\}$

$\omega_0 = (\alpha, \{\omega_1\}, \{\omega_0\})$

$\omega_1 = (\beta, \{\omega_1\}, \{\omega_1\})$

*Assume full communication. Then, following the first rule the revision yields* $t_1^c(\omega_0) = t_2^c(\omega_0) = \emptyset$ *and following the second rule, the revision yields* $t_2^c(\omega_0) = \{\omega_0\}$ *while* $t_1^c(\omega_0) = \emptyset$, *which is not possible in an MBS.*

The problem with the second rule exhibited in Example 7 reflects the contradiction between agent 1's initial beliefs and his interpretation of the other agent's announcement. Observe that the first revision rule has the same problem and cannot be applied here either. We reached now the difficult part in the construction of a general revision rule namely, the need to specify how agents *deal with contradictions* between their initial beliefs and the reported beliefs of the other agents.

## 4.2 Coping with contradictions

The two revision rules introduced above are formally not applicable when contradictions occur, that is, if there is no world among the ones initially believed by an agent that is compatible with the announcements of the other agents. However, the logic behind these two rules could be extended to deal with contradictions. Along the intuition of the first revision rule, the agent could adopt the beliefs announced by the other. In Example 7, agent 1's beliefs would now be given by $t_1^c(\omega_0) = \{\omega_0\}$. This corresponds to take at face value 2's announcements and, in particular, to admit that state $\alpha$ is true, something 1 did not believe in to begin with. Observe however that agent 1 is not proven wrong in his belief that the state of nature is $\beta$. Indeed, the only mistake that is revealed is that 1 believed 2 believed the state of nature was $\alpha$. We do not pursue this logic in the rest of the paper and concentrate on the logic behind the second revision rule: in presence of contradictions, the agent holds on to the beliefs that are not contradicted and changes in a "minimal way" the ones that are contradicted. Consider Example 7 again: what does agent 2 do when he hears 1's announcement that the state is $\alpha$, which is contradicting his initial beliefs? A plausible revised MBS, after full communication is:

$\omega_0 = (\alpha, \{\omega_1\}, \{\omega_0\})$
$\omega_1 = (\beta, \{\omega_1\}, \{\omega_0\})$

This revision is minimal in the sense that the initial disagreement on the state of nature persists. Agent 2 has only revised his beliefs by taking into account 1's beliefs, which explains why $t_2^c(\omega_1) = \{\omega_0\}$. The system is then closed by imposing that this minimal change becomes common belief. The general revision rule we'll introduce shortly is hence built on the idea that an agent keeps the beliefs that are not contradicted. This is illustrated in the next example.

**Example 8** *Let $S = \{\alpha, \beta, \gamma\}$ and $I = \{1, 2, 3\}$. Consider $\Omega = \{\omega_0, \omega_1, \omega_2\}$*

$\omega_0 = (\alpha, \{\omega_1\}, \{\omega_2\}, \{\omega_0\})$
$\omega_1 = (\beta, \{\omega_1\}, \{\omega_1\}, \{\omega_1\})$
$\omega_2 = (\gamma, \{\omega_1\}, \{\omega_2\}, \{\omega_2\})$

*Assume communication $I^c = \{2\}$. Agent 1 realizes that he was mistaken about 2's beliefs: 2 actually disagrees with 1 on both the state of nature and 3's beliefs. A plausible revision would be:*

$\omega_0 = (\alpha, \{\omega_1\}, \{\omega_2\}, \{\omega_0\})$
$\omega_1 = (\beta, \{\omega_1\}, \{\omega_2\}, \{\omega_1\})$
$\omega_2 = (\gamma, \{\omega_1\}, \{\omega_2\}, \{\omega_2\})$

*where 1 only modified his beliefs about 2's beliefs (and the latter are common belief).*

We are not yet done: the principles discussed so far are still not enough to yield a satisfactory revision rule, as can be seen on the following example.

**Example 9** *Let $S = \{\alpha, \beta, \gamma\}$ and $I = \{1, 2, 3\}$. Consider $\Omega = \{\omega_0, \omega_1\}$*

$\omega_0 = (\alpha, \{\omega_1, \omega_2\}, \{\omega_0\}, \{\omega_1\})$
$\omega_1 = (\beta, \{\omega_1, \omega_2\}, \{\omega_1, \omega_2\}, \{\omega_1\})$
$\omega_2 = (\gamma, \{\omega_1, \omega_2\}, \{\omega_1, \omega_2\}, \{\omega_2\})$

*Assume communication $I^c = \{2, 3\}$. Agent 1 realizes he was mistaken about 2's beliefs but not necessarily about 3's beliefs. Thus, he will not necessarily keep all his initial beliefs $\{\omega_1, \omega_2\}$, and for instance, he might abandon $\omega_2$, yielding the following revised MBS:*

$\omega_0 = (\alpha, \{\omega_1\}, \{\omega_0\}, \{\omega_1\})$
$\omega_1 = (\beta, \{\omega_1\}, \{\omega_0\}, \{\omega_1\})$

To capture the phenomenon at work in Example 9, we need to add a sort of *personal attitude* of the agents as part of the data of the model. To capture this personal attitude of the agent, we assume that given an MBS, $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ and a communication $I^c$, there is an *order*, $\succeq_i^c$ for each $i$, which is a complete and transitive binary relation defined on $\Omega_i$, the set of states that $i$ could believe (recall that $\Omega_i = \cup_{\omega\in\Omega} t_i(\omega)$). For $\omega, \omega' \in \Omega_i$, $\omega \succeq_i^c \omega'$ is interpreted as saying that, given the communication $I^c$, agent $i$ believes that $\omega$ is "closer" to the true (unknown) world than $\omega'$ is. In addition to completeness and transitivity we shall assume that the order $\succeq_i^c$ is *consistent*:

**Definition 7** *Given an MBS,* $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ *and a communication* $I^c$*, an order* $\succeq_i^c$ *is said to be* consistent *if whenever* $\Omega_i \cap (\{\omega \in \Omega | t_j(\omega) = t_j(\omega_0), \forall j \in I^c\}) \neq \emptyset$,

$$\left[\omega \in \Omega_i \cap \{\omega'' \in \Omega | t_j(\omega'') = t_j(\omega_0), \forall j \in I^c\}\right] \Leftrightarrow \omega \succeq_i^c \omega'; \quad \forall \omega' \in \Omega_i$$

Thus, an order $\succeq_i^c$ is consistent if it ranks highest every state of the world that is initially deemed to be possibly believed by $i$ (i.e., states that are in $\Omega_i$) and that explains (is compatible with) the others' announcements. This requirement is really rather weak: in Example 9, it does not impose anything on $\succeq_1^c$ and the three different orders $\omega_1 \succ_1^c \omega_2$, $\omega_2 \succ_1^c \omega_1$, or $\omega_1 \sim_1^c \omega_2$ are all consistent.

These orders could be used to model various assumptions about the expertise of the different agents. Consider for instance the case in which all agents are equally competent. An order representing this is to say that the different states are ordered according to the number of inconsistencies: if, for instance, two agents announce beliefs that are in contradiction with the one they should have in state $\omega'$, then this state is ranked lower than state $\omega''$ in which only one agent announces beliefs that are in contradiction. Another possible ranking, representing the idea that, say, agent 1 is known to be an expert, would be to rank states according to whether agent 1's announcement are in contradiction or not with the state.

The revision rule we are about to introduce is based on these orders and on the assumption that, loosely speaking, they are commonly believed by all agents so as to enable interactive reasoning about mutual beliefs. It should be noted that $i$'s order is defined on $\Omega_i$, which does not, in general, coincide with $i$'s belief horizon. This is important since $j$ might (mistakenly) believe that there are states that $i$ considers possible (i.e., states in $BH_j(\omega_0, t) \cap \Omega_i$). Hence, $j$ needs to know how to revise $i$'s beliefs in these worlds. Implicit in the fact that the order introduced for agent $i$ is defined on all of $\Omega_i$ is the idea that all agents agree on how to revise $i$'s beliefs. If states $\omega$ and $\omega'$ belong to both $j$ and $j'$'s belief horizon, then these two agents agree on which is ranked highest according to $i$'s ordering. Furthermore, this fact is commonly believed by all agents. Hence, we'll make the maintained assumption that given an MBS, $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ and a communication, all orders $\succeq_i^c$ for $i \in I$ are consistent and commonly known by all agents (in the sense we just discussed).

### 4.3 A general revision rule: definition and examples

We now propose a general revision rule that copes with announcements contradicting initial beliefs. We first define the rule and then illustrate it via a few examples. Given an MBS and a communication, the revision rule we propose consists of two elements.

- **Step 1** Each agent $i$ retains all states of the world in the set $\Omega_i$ that have the highest rank in his order $\succeq_i^c$.

- **Step 2** In the states retained, the beliefs attributed to other agents are constructed by taking into account the modifications they have made in step 1. This corresponds to the idea that the way agents modify their beliefs is common belief.

Note that step 2 is possible since by assumption the announcements $t_i(\omega_0)$, $i \in I^c$, and the orders $\succeq_i^c$ are " commonly known", hence the modification made can be "performed" by player $i$ for each player $j$.

**Definition 8** *Let $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ be a minimal MBS and consider the communication $I^c$.*[7] *The* revision *of $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ is $\left(\Omega^c, \omega_0, s, (t_i^c)_{i\in I}\right)$, defined in two steps:*

*First define $\tilde{t}_i(.)$ by $\forall \omega \in \Omega$, $\tilde{t}_i(\omega) = \{\omega' \in t_i(\omega) | \omega' \succeq_i^c \omega'', \ \forall \omega'' \in t_i(\omega)\}$ and let $\widetilde{\Omega} = \{\omega_0\} \cup \left(\cup_{i\in I} BH_i(\omega_0, \tilde{t})\right)$*

*Then, define $t_i^c(.)$ as follows:*

- $\forall \omega \in \Omega$, $\forall i \in I \setminus I^c$, $t_i^c(\omega) = \tilde{t}_i(\omega)$

- $\forall \omega \in \Omega$, $\forall i \in I^c$, $t_i^c(\omega) = \tilde{t}_i(\omega_0)$

*and set $\Omega^c = \{\omega_0\} \cup (\cup_{i\in I} BH_i(\omega_0, t^c))$*

**Remark 2** *As in the two previous rules, there is a slight abuse of notation in the previous definition, as $\widetilde{\Omega}$ is defined via belief horizons that are only defined once $\widetilde{\Omega}$ is given. The same problem arises for $\Omega^c$ and can be dealt with in the same manner as in Remark 1.*

The logic of the revision rule we propose can be understood as follows. The first step is to eliminate in one's beliefs the worlds that are considered the farthest away (after hearing the others' announcements) from the true world, according to the order $\succeq_i^c$. In this operation, the agent considers his initial beliefs as valid and simply gets rid of the states that are not ranked highest with respect to his ordering. Hence, as discussed in section 4.2, agents are assumed to anchor their revised beliefs in their initial beliefs. This step of the revision procedure coincides with the second revision rule when it is well defined and the agents' orders are consistent.

[7] We define the revision rule only for minimal MBS, since otherwise the outcome of the revision process depends on the representation used, as can be seen in Example 13 in appendix B.

The second step of the revision procedure consists in dealing with the remaining "contradictions". One could interpret it as follows. After the first step of the revision procedure, agents announce their (corrected) beliefs: if, after the first step, the MBS obtained contains states that specify different beliefs for say agent $i$ than the ones he announces, then simply replace these beliefs by his announcement. This second step might be irrelevant if the announcements of the agents were all compatible with what they expected (this is the case under our maintained assumption for instance if the MBS is correct), in which case after the first step, all the states that were not ranked highest have been eliminated. We now illustrate this rule on Example 2.

**Example 10** *(Example 2 continued) Consider full communication and the following consistent orders for 1 and 2:*

$$\omega_3 \succ_1^c \omega_1 \succ_1^c \omega_2 \text{ and } \omega_1 \sim_2^c \omega_2 \succ_2^c \omega_3$$

*The first step of the definition yields the following MBS:*

$\omega_0 = (\alpha, \{\omega_1\}, \{\omega_3\})$
$\omega_1 = (\alpha, \{\omega_1\}, \{\omega_1, \omega_2\})$
$\omega_2 = (\beta, \{\omega_1\}, \{\omega_1, \omega_2\})$
$\omega_3 = (\beta, \{\omega_3\}, \{\omega_3\})$

*At the next step, the contradictions are treated by replacing with the announcement.*

$\omega_0 = (\alpha, \{\omega_1\}, \{\omega_3\})$
$\omega_1 = (\alpha, \{\omega_1\}, \{\omega_3\})$
$\omega_2 = (\beta, \{\omega_1\}, \{\omega_3\})$
$\omega_3 = (\beta, \{\omega_1\}, \{\omega_3\})$

*It is easy to check that in the above MBS, states $\omega_0$ and $\omega_1$ actually express the same hierarchy of beliefs, and that the same is true for states $\omega_2$ and $\omega_3$. Hence, it can be reduced (according to the formal process defined in Appendix A) to the following MBS:*

$\omega_0 = (\alpha, \{\omega_0\}, \{\omega_1\})$
$\omega_1 = (\beta, \{\omega_0\}, \{\omega_1\})$

*The outcome of the revision procedure is therefore a situation in which disagreement about the state of nature is common belief but becomes common belief.*

When the initial MBS is incorrect, the revision can lead to a modification of the beliefs so that they become correct. However, if agents disagree about the state of nature, that is if initially they believed in disjoint sets of states of nature, then the revision will never lead them to agree on the true state of nature and disagreement will persist.

## 4.4 General revision rule: consistency properties

We now proceed to show a certain number of properties of the revision rule. The first proposition states that the revision rule yields an MBS at each of the two steps of Definition 8. Hence, contrary to the rules discussed in Section 4.1, revision is always possible and does not lead to ill-defined belief systems.

**Proposition 5** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be a minimal MBS. Then,* $\left(\widetilde{\Omega}, \omega_0, s, (\tilde{t}_i)_{i\in I}\right)$ *and* $(\Omega^c, \omega_0, s, (t_i^c)_{i\in I})$ *are MBS.*

The next Proposition establishes a link between the second revision rule and the general one. This formally shows that the logic behind the general revision rule is indeed, as argued above, the one present in the second revision rule.

**Proposition 6** *Assume that agents' orders are consistent. Then, when the second revision rule is applicable, it coincides with the first step of the general revision rule, while the second step is void.*

A direct corollary to this Proposition is that when the MBS is totally correct then the second step of the revision process is void (i.e., $\widetilde{\Omega} = \Omega^c$) if the agents' orderings over the state space are consistent. This proposition also establishes that when the MBS is totally correct (a sufficient condition for the first two rules introduced above to be well defined), then all the revision rules we have defined coincide.

### 4.5 General revision rule: agreement and consensus

In this section, we seek to characterize conditions under which the revision leads to different forms of agreements among agents. This requires making a detour *via* the definition and characterization of *common S-beliefs systems*, in which agents' beliefs about the state of nature are common belief.

For a given MBS, $(\Omega, \omega_0, s, (t_i)_{i\in I})$ define the S-belief to be the event

$$SB(\omega_0, t) = \{\omega \in \Omega | s(t_i(\omega)) = s(t_i(\omega_0)) \ \forall i \in I\}$$

The $S$-belief is the event "for all $i \in I$, agent $i$ believes that the *state of nature* is in $s(t_i(\omega_0))$". In other words, $SB(\omega_0, t)$ is the subset of $\Omega$ in which the first level beliefs about $S$ are as those in $\omega_0$, i.e., the beliefs in the true state. We define now a special case of belief system, where the first level beliefs about $S$ are common beliefs.

**Definition 9** *An MBS,* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *is a* common S-belief system (*henceforth CSBS*) *if* $SB(\omega_0, t)$ *is common belief.*

In a CSBS, the agents' beliefs about the state of nature are common beliefs. Agents need not agree in a CSBS. It is thus possible to represent situations in which agents' disagreement is common belief. Example 4 is an instance of such a situation: 1 believes $\alpha$, 2 believes $\alpha$ or $\beta$ and this is common belief, i.e., agents disagree and this disagreement is common belief. We now establish properties about the degree to which agents agree after communication and revision have occurred. When *all* agents communicate, the revision leads to a situation in which beliefs about the state of nature are common belief. When agents still disagree about the state of nature, this models situation in which this disagreement is common belief. Such a case is illustrated in Example 10 above.

**Proposition 7** *Let* $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ *be a minimal MBS. Then,* $\left(\Omega^c, \omega_0, s, (t_i^c)_{i\in I}\right)$ *is a CSBS whenever* $I^c = I$.

When the initial MBS is already a CSBS, that is, when the beliefs about the state of nature of all agents are common belief, then communication does not lead to any further revision.

**Proposition 8** *Let* $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ *be a minimal CSBS. Then, if agents' orderings* $\succeq_i^c$ *are consistent,* $\left(\Omega^c, \omega_0, s, (t_i^c)_{i\in I}\right) = \left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$

The notion of CSBS does not entail a strong notion of agreement since indeed, disagreement can be common belief. A particular case of a CSBS is when the first level beliefs of all agents are the same, and thus $t_i(\omega) = t_j(\omega)$ for all $i, j \in I$ and all $\omega \in \Omega$. This represents a situation of consensus, when all agents have the same beliefs.

**Definition 10** *A minimal MBS,* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *is* consensual *if for all* $i, j \in I$, $t_i(\omega_0) = t_j(\omega_0)$.[8]

We now give a sufficient condition that entails that revision leads to a consensual MBS.

**Proposition 9** *Let* $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ *be a minimal MBS and assume it is totally correct. Assume further that* $I^c = I$ *and that agents' orderings* $\succeq_i^c$ *are consistent. Then,* $\left(\Omega^c, \omega_0, s, (t_i^c)_{i\in I}\right)$ *is consensual.*

This Proposition establishes that only under rather strong assumption will the revision process lead to a consensual belief system, in which all agents agree. Indeed, the assumption that the MBS be totally correct is necessary to get consensus, as can be seen on Example 4, in which the MBS is correct but not totally correct and no revision occurs after full communication.

## 4.6 General revision rule: dynamics

We now extend the static framework considered so far to study situations in which announcements are made sequentially. A communication sequence of length $T$, is the specification of a sequence of sets $\{I_\tau^c\}_{\tau=1,\dots,T}$ and of announcements at each stage $\{t_{i,\tau}(\omega_0)\}_{i\in I_\tau^c, \tau=1,\dots T}$ since here also we'll restrict attention to communication sequences in which agents announce precisely their true beliefs. We'll say that the communication is *exhaustive* if $\cup_{\tau=1,\dots,T} I_\tau^c = I$, i.e., if all agents announce at some point in time. One can also easily adapt the definition of the orders to take into account this temporal aspect (it is enough to have orders indexed by $\tau$) as well as the notion of consistency (which must hold at each given date).

[8] Recall that the MBS we are interested in are minimal. If the MBS is not minimal then the definition of consensus need to be modified: an MBS is said to be consensual if it has a representation that is consensual.

The sequential rule of revision in that case is a straightforward extension of the revision rule proposed in Definition 8. This rule is implemented at each stage, yielding an MBS at stage $\tau$ denoted $\Omega^c_\tau$. Recall however that, without further restrictions on agents' ordering of the states, the revision rule has to be applied to MBS that are minimal. Hence, if at the end of any given stage, the resulting MBS is not minimal, then we replace it by one of its minimal representations (as defined in appendix A) before proceeding to the next round of announcement/revision. In this process, we always make sure that the labelling of the true state remains $\omega_0$ at all stages. The revision process is well defined in the sense that it does not depend on the choice of the minimal representation (see Proposition 15 in appendix B).

Of particular interest in this dynamic setting are first whether agreement is eventually reached and second, whether the order of the announcements (who announces when) might matter for the situation eventually reached. We answer these two questions affirmatively.

**Proposition 10** *Let* $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ *be a minimal MBS and assume the communication* $(I^c_\tau)_{\tau=1,\dots,T}$ *is exhaustive, then* $\left(\Omega^c_T, \omega_0, s, \left(t^c_{i,T}\right)_{i\in I}\right)$ *is a CSBS.*

The revision process ends when the smallest $k$ such that $\cup_{\tau=1,\dots,k} I^c_\tau = I$ is reached. Hence, we established that convergence occurs and at the point of convergence, beliefs about the state of nature are common beliefs (but might be different). The next point we address is whether the order of announcements matters and show that it does not if the MBS is totally correct, but might otherwise.

**Proposition 11** *Let* $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ *be a minimal totally correct MBS. Consider two sequential communications* $(I^c_\tau)_{\tau=1,\dots,T}$ *and* $(\bar{I}^c_\tau)_{\tau=1,\dots,\bar{T}}$ *of length* $T$ *and* $\bar{T}$ *respectively, such that* $\cup_{\tau=1,\dots,T} I^c_\tau = \cup_{\tau=1,\dots,\bar{T}} \bar{I}^c_\tau$*. Assume finally that agents have consistent orders at any point in time. Then, the revision rule leads to two equivalent MBS.*

The Proposition provides a rather strong sufficient condition (that the MBS is totally correct) under which the order of announcement does not matter. This sufficient condition can be relaxed but not much. In Example 11, it is shown that as soon as one has to cope with contradictions, the order matters. One may wonder whether commutativity holds when the second rule is applicable, i.e., when there is no contradiction. Example 12 is a case of a correct MBS in which the second revision rule is applicable when all agents announce simultaneously but is not for a sequential announcement. In this case, the order does matter. This points out the fact that whether an agent will have to deal with contradiction depends on the order of announcements. We conjecture that for two sequential communication with the same set of agents announcing their beliefs, the two revised MBS will be the same whenever the second rule is applicable.

**Example 11** *Let* $S = \{\alpha, \beta, \gamma\}$ *and* $I = \{1, 2, 3\}$*. Consider* $\Omega = \{\omega_0, \omega_1, \omega_2\}$ *where*
$\omega_0 = (\alpha, \{\omega_1, \omega_2\}, \{\omega_0, \omega_1\}, \{\omega_0, \omega_2\})$

$\omega_1 = (\beta, \{\omega_1, \omega_2\}, \{\omega_0, \omega_1\}, \{\omega_1\})$
$\omega_2 = (\gamma, \{\omega_1, \omega_2\}, \{\omega_2\}, \{\omega_0, \omega_2\})$

- *First consider the case where there is only one round of announcement and* $I^c = \{1, 2, 3\}$. *Given this announcement, it has to be the case that any consistent order has that* $\omega_0 \succ_i^c \omega_2$ *and* $\omega_0 \succ_i^c \omega_1$ *for* $i = 2, 3$. *Observe furthermore that only agent 1 is faced with a contradiction: he did not contemplate any state of the world in which he was expecting agent 2 and 3's simultaneous announcement. The only three consistent orders that are possible for agent 1 are therefore:* $\omega_1 \sim_1^c \omega_2$, *or* $\omega_1 \succ_1^c \omega_2$, *or* $\omega_2 \succ_1^c \omega_1$. *We give the outcome of the revision in these three cases:*
  - *If* $\omega_1 \sim_1^c \omega_2$, *then the revision rule yields the following MBS*
    $\omega_0 = (\alpha, \{\omega_1, \omega_2\}, \{\omega_0\}, \{\omega_0\})$
    $\omega_1 = (\beta, \{\omega_1, \omega_2\}, \{\omega_0\}, \{\omega_0\})$
    $\omega_2 = (\gamma, \{\omega_1, \omega_2\}, \{\omega_0\}, \{\omega_0\})$
  - *If* $\omega_1 \succ_1^c \omega_2$, *then the revision rule yields the following MBS*
    $\omega_0 = (\alpha, \{\omega_1\}, \{\omega_0\}, \{\omega_0\})$
    $\omega_1 = (\beta, \{\omega_1\}, \{\omega_0\}, \{\omega_0\})$
  - *If* $\omega_2 \succ_1^c \omega_1$, *then the revision rule yields the following MBS*
    $\omega_0 = (\alpha, \{\omega_2\}, \{\omega_0\}, \{\omega_0\})$
    $\omega_2 = (\gamma, \{\omega_2\}, \{\omega_0\}, \{\omega_0\})$

- *Consider now the case where 1 and 2 announce first, revision occurs, and then 3 announces, that is,* $I_1^c = \{1, 2\}$ *and* $I_2^c = \{3\}$. *In the first round, the only possible consistent orders are that* $\omega_1 \succ_1^c \omega_2$ *for agent 1 and* $\omega_0 \succ_i^c \omega_1$ *and* $\omega_0 \succ_i^c \omega_2$ *for* $i = 2, 3$. *Thus, the revised MBS after the first round is given by*

  $\omega_0 = (\alpha, \{\omega_1\}, \{\omega_0, \omega_1\}, \{\omega_0\})$

  $\omega_1 = (\beta, \{\omega_1\}, \{\omega_0, \omega_1\}, \{\omega_1\})$

  *The same type of computation after 3's announcement yields:*

  $\omega_0 = (\alpha, \{\omega_1\}, \{\omega_0\}, \{\omega_0\})$

  $\omega_1 = (\beta, \{\omega_1\}, \{\omega_0\}, \{\omega_0\})$

- *Finally, consider the case where 1 and 3 announce first, revision occurs and then 2 announces, that is,* $I_1^c = \{1, 3\}$ *and* $I_2^c = \{2\}$. *Here again, consistent orders can be determined and we obtain after the first round:*

  $\omega_0 = (\alpha, \{\omega_2\}, \{\omega_0\}, \{\omega_0, \omega_2\})$

  $\omega_2 = (\gamma, \{\omega_2\}, \{\omega_2\}, \{\omega_0, \omega_2\})$

  *and finally we have, after 2's announcement:*

  $\omega_0 = (\alpha, \{\omega_2\}, \{\omega_0\}, \{\omega_0\})$

  $\omega_2 = (\gamma, \{\omega_2\}, \{\omega_0\}, \{\omega_0\})$

*Thus, we end up with different MBS according to the order of announcements.*

In this example, observe that non-commutativity does not come from possible inconsistencies in the orders. Non-commutativity comes from the fact that agents' revisions are done sequentially without keeping a memory of the reason why they changed their initial beliefs to begin with. This absence of memory explains why, in the sequential process in which 1 and 3 announce first and 2 second, 1 does not reconsider the elimination of $\omega_2$ (made upon 3's announcement) when 2 announces in the second stage. In the next example, the outcome of the revision process depends on the sequence of announcements although the MBS is initially correct (but not totally correct).

**Example 12** *Let $S = \{\alpha, \beta, \gamma\}$ and $I = \{1, 2, 3\}$. Consider $\Omega = \{\omega_0, \omega_1, \omega_2, \omega_3\}$ where*

$\omega_0 = (\alpha, \{\omega_0, \omega_1\}, \{\omega_0\}, \{\omega_0\})$
$\omega_1 = (\beta, \{\omega_0, \omega_1\}, \{\omega_0\}, \{\omega_2\})$
$\omega_2 = (\alpha, \{\omega_0, \omega_1\}, \{\omega_3\}, \{\omega_2\})$
$\omega_3 = (\gamma, \{\omega_3\}, \{\omega_3\}, \{\omega_3\})$

- *First consider the case where all agents announce simultaneously: $I^c = \{1, 2, 3\}$. The only consistent order for agent 1 must rank $\omega_0$ and $\omega_1$ in the following way: $\omega_0 \succ_1^c \omega_1$. Then we obtain the following MBS:*

  $\omega_0 = (\alpha, \{\omega_0\}, \{\omega_0\}, \{\omega_0\})$

- *Consider now the case where $I_1^c = \{1, 2\}$ and $I_2^c = \{3\}$. Observe that the "elimination stage" is irrelevant in the first revision, and the only operation to do is to replace any inconsistencies by the announcements of agent 1 and 2, yielding:*

  $\omega_0 = (\alpha, \{\omega_0, \omega_1\}, \{\omega_0\}, \{\omega_0\})$

  $\omega_1 = (\beta, \{\omega_0, \omega_1\}, \{\omega_0\}, \{\omega_2\})$

  $\omega_2 = (\alpha, \{\omega_0, \omega_1\}, \{\omega_0\}, \{\omega_2\})$

  *Note that $\omega_3$ has been dropped since, after the revision, it does not belong to any belief horizon. Before proceeding to the second round of revision, it is important to observe that the MBS after the first round is not minimal since state 1 and 2 are identical. Hence, it has a minimal representation:*

  $\omega_0 = (\alpha, \{\omega_0, \omega_1\}, \{\omega_0\}, \{\omega_0\})$

  $\omega_1 = (\beta, \{\omega_0, \omega_1\}, \{\omega_0\}, \{\omega_0\})$

  *Now, consider the second step, in which 3 announces his beliefs, i.e., $\omega_0$. This does not lead to any further revision.*

*Hence, the MBS we end up with is different from the one in which all agents were making their announcements simultaneously, showing that the order of these announcements matter, even though the initial MBS was correct.*

This last example has the feature that 1 believes that when the state of nature is $\beta$, 3 is mistaken about 2's beliefs. Hence, when 2 announces in the first round, 1 knows that 3's beliefs about 2 are now correct and no further revision takes place. On the other hand, when the three agents announce simultaneously, 3's announcement is enough to get rid of $\omega_2$, i.e., when 3 announces his beliefs, 1 learns that the state of nature is not $\beta$.

The two previous examples show that the revision process is not necessarily commutative, unless the initial MBS is well behaved (i.e., totally correct) as established in Proposition 11. This points out few interesting issues. First, the non commutativity is not directly linked to the procedure we adopted to treat announcements that are in contradiction with the initial beliefs of the agents. Indeed, in Example 12, the two sequential processes studied do not entail any contradiction: in both cases, the announcements made in the first round are compatible with part of the initial beliefs. Thus, agents only keep those states that are exactly compatible with the announcements. Second, non-commutativity of the revision procedure arises because agents treat each new MBS afresh, without keeping a memory of how they arrived at it. In that respect the sequential revision process we have described is "myopic". Another way of saying this is to describe the revision process we have defined as a markovian process: at each stage, the only information taken into account to revise is the state of the system at that stage. An alternative, more demanding, way of modelling things would be to go back, after each round of announcement, to the initial MBS and use all the sequence of announcements made up to that point in time to revise it. It is not clear whether the framework developed here is the most appropriate to treat this way of revising. Further, the "unbounded" memory assumption that this alternative approach would require might be too demanding in terms of the amount of information agents would have to keep at each stage of the revision process. Indeed, it is not necessary for totally correct MBS. Here again, an intuition that is correct in the absence of mistakes (i.e., the path through which one arrives at a given state of the epistemic system is not relevant) appears to be misleading in the more general case. Finally, non commutativity points out the fact that communication has another strategic aspect to it beyond its mere content: the order of the agenda (i.e., who gets to speak when) is important and agents are bound to take this into account if they have the choice as to when to speak (as is recognized in the "herd behavior" literature, see e.g., Gul and Lundholm (1995)).

Beyond these general remarks, we would like to argue that there is an important difference between the two examples of non-commutativity. In Example 11, non-commutativity is problematic: for instance, in the case where agents 1 and 2 announce first and then 3 announces, agent 1 should be allowed to reconsider the elimination of state $\omega_2$, since 2's and 3's announcements have essentially the same value to agent 1. The situation in Example 12 is different: the mistakes were not on the first level beliefs but on higher order beliefs. Hence, these beliefs do change after a first round of announcement. Thus, the non-commutativity of the rule is simply the reflect that higher order mistaken beliefs are corrected according to the announcements made at a given stage, before further revision is done.

# 5 Concluding Remarks

We have studied in this paper the representation of beliefs in a framework general enough to accommodate the presence of mistaken beliefs. The general setting developed was then used to study the revision of beliefs when agents are allowed to communicate truthfully their beliefs. At this stage of our work, we have focussed on communication processes in which agents announce truthfully and exactly their complete beliefs. We showed that the presence of mistakes can explain disagreement even after this form of communication (which forces consensus when there is no mistakes) has taken place. The communication process we considered could be interpreted as a process in which agents tell each other all what they believe. This is a rather natural notion from where to start. We could generalize the rule to communication in which agents do not necessarily announce exactly their beliefs but potentially some super set, but this does not seem to us the most promising avenue for further research. Rather, we believe that the next step is to come up with ways of representing more general "communication rules". Take for instance the three hat examples: there, the three girls answer a well defined question, "can you tell the color of your hat?". They do not communicate to each other the entire hierarchy of beliefs but only the answer to this question. How to represent such a communication rule in our framework is an open issue. Only after such a work has been done could we proceed to analyze convergence properties of communication processes that are less demanding than the one we studied here. Another interesting issue for future research is to allow for strategic communication among agents.

Another interesting issue is to model private information in our setting. More specifically, assume that after communication, agents agree and this is common belief. If they are then given some private information (from an external source), communicating their revised beliefs should reveal the private information they got. The situation might well be rather different when the MBS one starts from entails some (commonly believed) disagreement. Then, communicating the revised beliefs, after reception of (outside) private information might not be enough to fully reveal which information each agent got. Since revelation of private information is at the heart of no-trade theorems (see Milgrom and Stokey (1982)), we might conjecture that such theorems would not necessarily hold in this (extended) setting. Thus, it would be interesting to know whether (and when) the no trade result remains valid in a setting in which mistaken beliefs are allowed. In a similar vein, one could wonder, starting from a situation with mistaken beliefs, whether additional private information would lead to a correct belief structure (and restore no trade results).

## Appendix A: Equivalent Representations and Minimality of MBS

We introduce here the notion of equivalence of two beliefs systems. An MBS can be represented by another MBS if the entire belief hierarchy of the first is captured by the second. Formally, this means that there exists an onto relationship between the two MBS which projects two (or more) states of the first MBS onto a unique state in the second MBS. In that case, the two states in the first MBS, while formally different, are actually representing the same situation of mutual beliefs.

**Definition 11** *An MBS,* $(\Omega', \omega_0', s', (t_i')_{i\in I})$*, is a* representation *of the MBS,* $(\Omega, \omega_0, s, (t_i)_{i\in I})$*, if there exists a mapping* $\sigma$ *from* $\Omega$ *to* $\Omega'$ *such that*
*(i)* $\sigma(\Omega) = \Omega'$
*(ii)* $\sigma(\omega_0) = \omega_0'$
*(iii)* $s' \circ \sigma = s$
*(iv)* $\forall i \in I,\ t_i' \circ \sigma = \sigma \circ t_i$.

**Definition 12** *Two MBS,* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *and* $(\Omega', \omega_0', s', (t_i')_{i\in I})$*, are* equivalent *if they have a common representation, i.e., if there exists an MBS,* $(\Omega'', \omega_0'', s'', (t_i'')_{i\in I})$*, that is a representation of both* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *and* $(\Omega', \omega_0', s', (t_i')_{i\in I})$*.*

A special case of equivalence of two MBS is when each is a representation of the other, in which case the two spaces are identical up to a renaming of the states ($\sigma$ is hence a bijection). It is relatively easy to show that this notion of equivalence is in fact an equivalence relationship. Note that, by definition, the relation is symmetric. It is also reflexive since any MBS is equivalent to itself via the identity mapping. Transitivity can also be established (we do no report the proof for sake of brevity). We now define a notion of redundancy within an MBS.

**Definition 13** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be an MBS. Two states* $\omega_1, \omega_2 \in \Omega$ *are said to be* identical *if there exists an MBS,* $(\Omega', \omega_0', s', (t_i')_{i\in I})$ *and a mapping* $\sigma : \Omega \to \Omega'$ *as in Definition 11 such that* $\sigma(\omega_1) = \sigma(\omega_2)$*.*

Two states of the world are thus identical if there exists a representation of the MBS in which these two states are represented by the same state of the world. Our next step is to define minimal MBS, in which such a problem does not arise.

**Definition 14**

- *An MBS,* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *is* minimal *if no two distinct states of the world* $\omega, \omega' \in \Omega$*, are identical.*
- *An MBS,* $(\Omega', \omega_0', s', (t_i')_{i\in I})$ *is a* minimal representation *of* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *if it is a representation of* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *and it is minimal.*

**Proposition 12** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *be an MBS. Then it has a minimal representation* $(\Omega', \omega'_0, s', (t'_i)_{i\in I})$ *and all its minimal representations are equivalent.*

**Proposition 13** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *and* $(\Omega', \omega'_0, s', (t'_i)_{i\in I})$ *be minimal and equivalent MBS. Then there exists a one-to-one and onto mapping* $\phi$ *from* $\Omega$ *to* $\Omega'$ *such that conditions (i) to (iv) of definition 11 hold.*

In the paper we deal exclusively with minimal MBS. This is without loss of generality as we just saw that non minimal MBS always have a minimal representation.

## Appendix B: Minimality and Revision of MBS

In this appendix, we tackle the issue of whether the revision process we defined depend (in a meaningful way) on the representation of the MBS we consider. We first establish that if an MBS is correct so must be any representation of it.

**Proposition 14** *Let* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *and* $(\Omega', \omega'_0, s', (t'_i)_{i\in I})$ *be minimal and equivalent MBS. If* $\Omega$ *is correct then,* $\Omega'$ *is also correct.*

The following example illustrates why the revision rule we defined is restricted to minimal MBS: if it were not the case, the outcome of the revision process might depend on the representation adopted.

**Example 13** *Let* $S = \{\alpha, \beta\}$ *and* $I = \{1, 2\} = I^c$. *Consider* $\Omega = \{\omega_0, \omega_1, \omega_2\}$ *where*

$\omega_0 = (\alpha, \{\omega_0\}, \{\omega_1, \omega_2\})$
$\omega_1 = (\beta, \{\omega_1\}, \{\omega_1, \omega_2\})$
$\omega_2 = (\alpha, \{\omega_2\}, \{\omega_1, \omega_2\})$

*Assume agent 2's ordering is given by* $\omega_1 \sim_2 \omega_2$, *which is consistent. Then, the revision gives the following MBS:*

$\omega_0 = (\alpha, \{\omega_0\}, \{\omega_1, \omega_2\})$
$\omega_1 = (\beta, \{\omega_0\}, \{\omega_1, \omega_2\})$
$\omega_2 = (\alpha, \{\omega_0\}, \{\omega_1, \omega_2\})$

*which admits the following minimal representation:*

$\omega_0 = (\alpha, \{\omega_0\}, \{\omega_0, \omega_1\})$
$\omega_1 = (\beta, \{\omega_0\}, \{\omega_0, \omega_1\})$

*as state* $\omega_0$ *and* $\omega_2$ *are identical.*

*Now, observe that the initial MBS is not minimal, and admits the following minimal representation:*

$\omega_0 = (\alpha, \{\omega_0\}, \{\omega_0, \omega_1\})$
$\omega_1 = (\beta, \{\omega_1\}, \{\omega_0, \omega_1\})$

*If one considers this MBS, 2's order is given by* $\omega_0 \succ_2^c \omega_1$, *and hence revision yields*

$\omega_0 = (\alpha, \{\omega_0\}, \{\omega_0\})$

*which is clearly different from the one obtained above. Hence, it is important to restrict the application to our revision rule to minimal MBS.*

*The reason for which there is a discrepancy between the two revised MBS although they were equivalent to begin with is that the order $\succeq_2^c$ does not recognize the fact that $\omega_0$ and $\omega_2$ are identical. Thus, one way to cope with this difficulty is to apply the revision rule only on minimal MBS. Another way would be to assume that the agents' orderings of the states are such that $\omega \sim_i^c \omega'$ whenever $\omega$ and $\omega'$ are identical (i.e., if there exists a representation that projects these two states onto the same state).*

In the last proposition of this appendix, we show that the sequential revision process does not depend on the choice of a minimal representation at each stage. We first need to define a notion of compatibility of an agent's ordering between two equivalent minimal MBS.

**Definition 15** *Let $(\Omega, \omega_0, s, (t_i)_{i\in I})$ and $(\Omega', \omega_0', s', (t_i')_{i\in I})$ be two minimal and equivalent MBS such that there exists a one-to-one and onto mapping $\phi : \Omega \to \Omega'$ as in Proposition 13. $\succeq^c$ and $\succeq^{c'}$ are* compatible *if, for all $\omega_1, \omega_2 \in \Omega$, $\omega_1 \succeq^c \omega_2$ if and only if $\phi(\omega_1) \succeq^{c'} \phi(\omega_2)$.*

**Proposition 15** *Let $(\Omega, \omega_0, s, (t_i)_{i\in I})$ and $(\Omega', \omega_0', s', (t_i')_{i\in I})$ be two minimal equivalent MBS. Assume that agents' orders are consistent and compatible.*
*Then $(\Omega^c, \omega_0, s, (t_i^c)_{i\in I})$ and $((\Omega')^c, \omega_0', (s'), ((t_i')^c)_{i\in I})$ are equivalent MBS.*

## Appendix C: Proofs

**Proof. [Proposition 1]** Assume $(v)$ and define the set $\Omega' \subset \Omega$ by

$$\Omega' = \{\omega_0\} \cup \left\{\omega \in \Omega | \exists r \in \mathbb{N} \text{ and } \{i_k\}_{k=1}^{k=r}, i_k \in I, i_r = i \text{ s.th. } \omega \in t_{i_1}(t_{i_2}(\ldots(t_{i_r}(\omega_0)))\right\}$$

We show that $(\Omega', \omega_0, s|_{\Omega'}, (t_i|_{\Omega'})_{i\in I})$ satisfies conditions $(i)$ to $(iv)$ of Definition 1. Conditions $(i)$, $(iii)$, and $(iv)$ are obvious. Consider $i \in I$, $\omega \in \Omega'$ and take $\omega' \in t_i|_{\Omega'}(\omega) = t_i(\omega)$. It is easy to see that by definition of $\Omega'$, $\omega' \in \Omega'$ which proves that $t_i|_{\Omega'}$ is a mapping from $\Omega'$ to $2^{\Omega'}$. Therefore, condition $(v)$ implies that $\Omega' = \Omega$ and thus condition $(v')$ holds.
Assume now $(v')$ and suppose there exists $\Omega' \subsetneq \Omega$ such that $(\Omega', \omega_0, s|_{\Omega'}, (t_i|_{\Omega'})_{i\in I})$ satisfy conditions $(i)$ to $(iv)$ of Definition 1. Hence, $\exists \omega \in \Omega \setminus \Omega'$. However by $(v')$, $\exists r \in \mathbb{N}$ and $\{i_k\}_{k=1}^{k=r}, i_k \in I, i_r = i$ s.th. $\omega \in t_{i_1}(t_{i_2}(\ldots(t_{i_r}(\omega_0)))$. Since $\omega_0 \in \Omega'$, then $t_{i_r}|_{\Omega'}(\omega_0) = t_{i_r}(\omega_0) \subset \Omega'$ since condition $(ii)$ applies. By induction, we can show that for all $k = 1, ..r$,

$$(t_{i_k}|_{\Omega'}(\ldots(t_{i_r}|_{\Omega'}(\omega_0))) = (t_{i_k}(\ldots(t_{i_r}(\omega_0))) \subset \Omega'$$

and thus $\omega \in \Omega'$ yielding a contradiction. ∎

**Proof. [Proposition 2]** For $i \in I$ consider

$$NH_i(\omega_0, t) = \left\{\omega \in BH_i(\omega_0, t) | \forall r \in \mathbb{N} \text{ and } \{i_k\}_{k=1}^{k=r}, i_k \in I, i_r = i \text{ s.th. } \omega \notin t_{i_1}(t_{i_2}(\dots(t_{i_r}(\omega_0)))\right\}$$

and suppose $NH_i(\omega_0, t) \neq \emptyset$. Consider

$$Y = BH_i(\omega_0, t) \backslash NH_i(\omega_0, t)$$

Note that $Y$ is strictly included in $BH_i(\omega_0, t)$ since $NH_i(\omega_0, t) \neq \emptyset$. Remark that trivially $t_i(\omega_0) \subset Y$ which shows that $Y \neq \emptyset$ and condition $(i)$ of Definition 2 is satisfied.

Consider $\omega' \in Y$ and $j \in I$. Since $\omega' \in BH_i(\omega_0, t)$, $t_j(\omega') \subset BH_i(\omega_0, t)$. Suppose that $t_j(\omega') \nsubseteq Y$ and thus there exists $\omega \in NH_i(\omega_0, t) \cap t_j(\omega')$. Since $\omega' \in Y$, there exists a sequence $\{i_k\}_{k=1}^{k=r}, i_k \in I$, $i_r = i$ such that $\omega' \in t_{i_1}(t_{i_2}(\dots(t_{i_r}(\omega_0)))$. Then define the sequence $\{i'_k\}_{k=1}^{k=r+1}$ by $i'_1 = j$, $i'_k = i_{k-1}$ for all $k = 2, .., r+1$. Note that $i'_{r+1} = i$. Then we have that $\omega \in t_{i'_1}(t_{i'_2}(\dots(t_{i'_{r+1}}(\omega_0)))$ which is a contradiction with $\omega \in NH_i(\omega_0, t)$. Thus, condition (ii) of Definition 2 is also satisfied. That proves that $BH_i(\omega_0, t)$ is not the minimal subset which satisfies these conditions.

Thus $NH_i(\omega_0, t) = \emptyset$ and

$$BH_i(\omega_0, t) \subset \left\{\omega \in \Omega | \exists r \in \mathbb{N} \text{ and } \{i_k\}_{k=1}^{k=r}, i_k \in I, i_r = i \text{ s.th. } \omega \in t_{i_1}(t_{i_2}(\dots(t_{i_r}(\omega_0)))\right\}$$

Conversely, consider $\omega \in \Omega$ such that there exists[9] $r \in \mathbb{N}$ and $\{i_k\}_{k=1}^{k=r}$, $i_k \in I$, $i_r = i$ such that $\omega \in t_{i_1}(t_{i_2}(\dots(t_{i_r}(\omega_0)))$ and let us suppose that $\omega \notin BH_i(\omega_0, t)$. Then there exists $\{\omega_k\}_{k=1}^{k=r}$ such that $\omega_r = \omega_0$, $\forall k = 1, .., r-1$ $\omega_k \in t_{i_{k+1}}(\omega_{k+1})$ and $\omega \in t_{i_1}(\omega_1)$. Since $\omega \notin BH_i(\omega_0, t)$, condition (ii) of Definition 2 implies that $\omega_1 \notin BH_i(\omega_0, t)$. Recursively, we have that $\forall k = 1, .., r-1$, $\omega_k \notin BH_i(\omega_0, t)$. Hence, since $\omega_{r-1} \notin BH_i(\omega_0, t)$, $t_i(\omega_0) \nsubseteq BH_i(\omega_0, t)$, contradicting $(i)$ of Definition 2. ■

**Proof. [Proposition 4]** Let $(\Omega, \omega_0, s, (t_i)_{i \in I})$ be a totally correct MBS and consider a communication $I^c$. We first check that for the two rules, the revision yields a well defined MBS. Remark that the collection $\left(\Omega^c, \omega_0, s, (t_i^c)_{i \in I}\right)$ satisfies conditions $(i)$ and $(iii)$ to $(v)$ of Definition 1. For condition $(ii)$, we have to check that $\forall \omega \in \Omega^c$, $\forall i \in I$, $t_i^c(\omega) \neq \emptyset$. Given Proposition 2, it is equivalent to show that for all sequence $\{i_k\}_{k=1}^{k=r}$, for all $\omega \in t_{i_1}^c(t_{i_2}^c(\dots(t_{i_r}^c(\omega_0)))$ and for all $i$, $t_i^c(\omega) \neq \emptyset$. Consider $\omega \in t_{i_1}^c(t_{i_2}^c(\dots(t_{i_r}^c(\omega_0)))$. There exists $\omega' \in \Omega^c$ such that $\omega \in t_{i_1}^c(\omega')$. For the first rule, that means that $t_{i_1}^c(\omega') = t_{i_1}(\omega') \cap (\cap_{j \in I^c} t_j(\omega_0))$ and thus $\omega \in (\cap_{j \in I^c} t_j(\omega_0))$. Yet, since $(\Omega, \omega_0, s, (t_i)_{i \in I})$ is a totally correct MBS, $\omega \in t_i(\omega)$ and therefore $\omega \in t_{i_1}^c(\omega)$ which proves that $t_i^c(\omega) \neq \emptyset$. For the second rule, that means that $t_{i_1}^c(\omega') = t_{i_1}(\omega') \cap \{\omega'' \in \Omega | t_j(\omega'') = t_j(\omega_0);\ \forall j \in I^c\}$ and thus $\omega \in \{\omega'' \in \Omega | t_j(\omega'') = t_j(\omega_0);\ \forall j \in I^c\}$ and therefore $t_i^c(\omega) \neq \emptyset$.

[9] By Proposition 1, such an $r$ exists

We now proceed to show that the second revision rule coincides with the first revision rule. To that end, it is enough to show that, for all $j \in I^c$, $\omega \in t_j(\omega_0)$ if and only if $t_j(\omega) = t_j(\omega_0)$. This is a direct consequence of the fact that the MBS is totally correct. Hence, the two definitions coincide when the MBS is totally correct. ■

**Proof. [Proposition 5]** It is straightforward to check conditions $(i)$ to $(v)$ of Definition 1 for both systems. ■

**Proof. [Proposition 6]** Let $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ be a minimal MBS and a communication $I^c$. Let suppose that the second revision rule is applicable and note $\left(\Omega^{c(2)}, \omega_0, s, (t_i^{c(2)})_{i\in I}\right)$ the resulting MBS. Denote $\widetilde{\Omega}$, $\left(\tilde{t}_i\right)_{i\in I}$ and $\left(\Omega^c, \omega_0, s, (t_i^c)_{i\in I}\right)$ as defined in Definition 8. Since the second revision rule is applicable, for all sequence $\{i_k\}_{k=1}^{k=r}$, for all $\omega \in t_{i_1}^{c(2)}(t_{i_2}^{c(2)}(\dots(t_{i_r}^{c(2)}(\omega_0)))$, and for all $i$, $t_i^{c(2)}(\omega) \neq \emptyset$.

For $i$ and $\omega_0$ we have that $t_i^{c(2)}(\omega_0) = \{\omega \in t_i(\omega_0) | t_j\left(\omega\right) = t_j(\omega_0);\ \forall j \in I^c\}$ and $\tilde{t}_i(\omega_0) = \{\omega' \in t_i(\omega) | \omega' \succeq_i^c \omega^*,\ \forall \omega^* \in t_i(\omega)\}$. Let $\omega \in t_i^{c(2)}(\omega_0)$. Then, $\omega \in \Omega_i \cap \{\omega' \in \Omega | t_j(\omega') = t_j(\omega_0), \forall j \in I^c\}$ and thus, since $\succeq_i^c$ is consistent, that implies $\omega \succeq_i^c \omega'$ for all $\omega' \in t_i(\omega_0)$ and thus $\omega \in \tilde{t}_i(\omega_0)$. Therefore, $t_i^{c(2)}(\omega_0) \subseteq \tilde{t}_i(\omega_0)$. Conversely, consider $\omega \in \tilde{t}_i(\omega_0)$. Since $t_i^{c(2)}(\omega_0) \subseteq \Omega_i \cap \{\omega' \in \Omega | t_j(\omega') = t_j(\omega_0), \forall j \in I^c\}$, thus $\Omega_i \cap \{\omega' \in \Omega | t_j(\omega') = t_j(\omega_0), \forall j \in I^c\} \neq \emptyset$ and since $\succeq_i^c$ is consistent, $\omega \in \Omega_i \cap \{\omega' \in \Omega | t_j(\omega') = t_j(\omega_0), \forall j \in I^c\}$. Therefore, $\omega \in t_i^{c(2)}(\omega_0)$ and thus $t_i^{c(2)}(\omega_0) = \tilde{t}_i(\omega_0)$. Furthermore, it is straightforward to check that $t_i^c(\omega_0) = \tilde{t}_i(\omega_0)$.

Let suppose that for $r \geq 1$, for all $r'$, such that $r \geq r' \geq 1$, for all sequence $\{i_k\}_{k=1}^{k=r'}$, for all $\omega \in t_{i_1}^{c(2)}(t_{i_2}^{c(2)}(\dots(t_{i_{r'}}^{c(2)}(\omega_0)))$, $t_i^{c(2)}(\omega) = t_i^c(\omega) = \tilde{t}_i(\omega)$. Consider a sequence $\{i_k\}_{k=1}^{k=r+1}$, and $\omega \in t_{i_1}^{c(2)}(t_{i_2}^{c(2)}(\dots(t_{i_{r+1}}^{c(2)}(\omega_0)))$. By assumption,

$$t_{i_1}^{c(2)}(t_{i_2}^{c(2)}(\dots(t_{i_{r+1}}^{c(2)}(\omega_0)))) = t_{i_1}^c(t_{i_2}^c(\dots(t_{i_{r+1}}^c(\omega_0)))) = \tilde{t}_{i_1}(\tilde{t}_{i_2}(\dots(\tilde{t}_{i_{r+1}}(\omega_0))))$$

and thus $\omega \in \tilde{t}_{i_1}(\tilde{t}_{i_2}(\dots(\tilde{t}_{i_{r+1}}(\omega_0))))$. Since the second revision rule is applicable,

$$t_i^{c(2)}(\omega) = t_i(\omega) \cap \left\{\omega' \in \Omega | t_j\left(\omega'\right) = t_j(\omega_0);\ \forall j \in I^c\right\} \neq \emptyset$$

and therefore, by a same argument as before, $t_i^{c(2)}(\omega) = \tilde{t}_i(\omega)$. If $i \in I \backslash I^c$, $t_i^c(\omega) = \tilde{t}_i(\omega)$. If $i \in I \setminus I^c$, $t_i^c(\omega) = \tilde{t}_i(\omega_0)$. Now, there exists $\omega' \in t_{i_2}^{c(2)}(\dots(t_{i_{r+1}}^{c(2)}(\omega_0)))$ such that

$$\omega \in t_{i_1}^{c(2)}(\omega') = t_{i_1}(\omega) \cap \left\{\omega'' \in \Omega | t_j\left(\omega''\right) = t_j(\omega_0);\ \forall j \in I^c\right\}$$

Therefore, $t_i\left(\omega\right) = t_i(\omega_0)$ and thus $\tilde{t}_i\left(\omega\right) = \tilde{t}_i(\omega_0) = t_i^c(\omega)$.

Thus we proved that for all sequence $\{i_k\}_{k=1}^{k=r}$, for all $\omega \in t_{i_1}^{c(2)}(t_{i_2}^{c(2)}(\dots(t_i^{c(2)}(\omega_0)))$, $t_i^{c(2)}(\omega) = t_i^c(\omega) = \tilde{t}_i(\omega)$. Therefore, $\Omega^{c(2)} = \widetilde{\Omega} = \Omega^c$ which completes the proof. ■

**Proof. [Proposition 7]**

Before proceeding to the proof of the Proposition itself, we need a lemma in which CSBS is characterized by the fact that any given agent must have the same beliefs in all the states of the world.

**Lemma 1** *Let $(\Omega, \omega_0, s, (t_i)_{i\in I})$ be a minimal MBS. Then, the following assertions are equivalent*

*(i)* $(\Omega, \omega_0, s, (t_i)_{i\in I})$ *is CSBS*

*(ii)* $SB(\omega_0, t) = \Omega$

*(iii)* $\forall \omega \in \Omega, \forall i \in I, \quad t_i(\omega) = t_i(\omega_0)$

**Proof.** [**Lemma 1**] We first prove $(i) \Leftrightarrow (ii)$. Since $SB(\omega_0, t)$ is common beliefs, we know by Corollary 2 that $\cup_{i\in I} BH_i(\omega_0, t) \subset SB(\omega_0, t) \subset \Omega = \{\omega_0\} \cup_{i\in I} BH_i(\omega_0, t)$. Note that by definition, $\omega_0 \in SB(\omega_0, t)$ and thus $SB(\omega_0, t) = \Omega$. Conversely, if $SB(\omega_0, t) = \Omega$, then $(\Omega, \omega_0, s, (t_i)_{i\in I})$ is a CSBS.

We next prove $(i) \Leftrightarrow (iii)$. From what we just proved, one way is obvious: since the condition $t_i(\omega) = t_i(\omega_0)$ $\forall \omega \in \Omega, \forall i \in I$ implies that $SB(\omega_0, t) = \Omega$, and hence the MBS is CSBS.

Conversely, assume that the MBS is CSBS. Then $SB(\omega_0, t) = \Omega$. Consider $\Omega' = \{\omega'_s\}_{s\in s(\Omega)}$ and the MBS, $\left(\Omega', \omega'_{s(\omega_0)}, s', (t'_i)_{i\in I}\right)$ defined by $\forall \omega'_s \in \Omega'$, $s'(\omega'_s) = s$ and $\forall i \in I$, $t'_i(\omega'_s) = \{\omega'_{s'} \in \Omega | s' \in s(t_i(\omega_0))\}$. Define the mapping $\sigma : \Omega \to \Omega'$ by $\forall \omega \in \Omega$, $\sigma(\omega) = \omega'_{s(\omega)}$. By construction, we have that $\sigma(\Omega) = \Omega'$, $\sigma(\omega_0) = \omega'_{s(\omega_0)}$, and $s' \circ \sigma = s$. Consider now $i \in I$ and $\omega \in \Omega$. Then

$$t'_i \circ \sigma(\omega) = t'_i\left(\omega'_{s(\omega)}\right) = \{\omega'_{s'} \in \Omega | s' \in s(t_i(\omega_0))\}$$

while

$$\begin{aligned} \sigma \circ t_i(\omega) &= \{\omega'_{s'} \in \Omega | \exists \omega'' \in t_i(\omega) \text{ such that } \omega'_{s'} = \sigma\left(\omega''\right)\} \\ &= \{\omega'_{s'} \in \Omega | \exists \omega'' \in t_i(\omega) \text{ such that } s' = s\left(\omega''\right)\} \\ &= \{\omega'_{s'} \in \Omega | s' \in s(t_i(\omega))\} \end{aligned}$$

But since $SB(\omega_0, t) = \{\omega \in \Omega | s(t_i(\omega)) = s(t_i(\omega_0)) \ \forall i \in I\} = \Omega$, we have

$$\sigma \circ t_i(\omega) = \{\omega'_{s'} \in \Omega | s' \in s(t_i(\omega_0))\} = t'_i \circ \sigma(\omega)$$

Thus $t'_i \circ \sigma = \sigma \circ t_i$ which shows that the MBS, $\left(\Omega', \omega_{s(\omega_0)}, s', (t'_i)_{i\in I}\right)$ is a representation of the MBS, $(\Omega, \omega_0, s, (t_i)_{i\in I})$. Since $(\Omega, \omega_0, s, (t_i)_{i\in I})$ is minimal, $\sigma$ is a one-to-one mapping. Remark now that by construction $\forall \omega'_s \in \Omega', \forall i \in I$, $t_i(\omega'_s) = t_i(\omega'_{s(\omega_0)})$ and since $\sigma^{-1}$ is a one-to-one mapping, $t_i(\sigma^{-1}(\omega'_s)) = t_i(\sigma^{-1}(\omega'_{s(\omega_0)}))$, establishing that $\forall \omega \in \Omega, \forall i \in I, \quad t_i(\omega) = t_i(\omega_0)$. ■

The proof of Proposition 7 is now trivial: If $I^c = I$, then by the construction of $t_i^c$ given in Definition 8, $\forall i$, $\forall \omega \in \Omega^c$, $t_i^c(\omega) = \tilde{t}_i(\omega_0)$ and thus according to Proposition 1, $\left(\Omega^c, \omega_0, s, (t_i^c)_{i\in I}\right)$ is a CSBS. ■

**Proof.** [**Proposition 8**] Since $\left(\Omega, \omega_0, s, (t_i)_{i\in I}\right)$ is a CSBS, Lemma 1 yields that $\forall \omega \in \Omega, \forall i \in I, \quad t_i(\omega) = t_i(\omega_0)$. Hence, $\Omega = \{\omega \in \Omega | t_j(\omega) = t_j(\omega_0), \forall j \in I^c\}$ and since agents' orderings $\succeq_i^c$ are consistent,

$$\forall \omega \in \Omega, \forall i \in I, \quad \tilde{t}_i(\omega) = t_i(\omega) = t_i(\omega_0)$$

Therefore, we also have $\forall \omega \in \Omega$, $\forall i$, $t_i^c(\omega) = t_i(\omega_0)$ and thus

$$\Omega^c = \{\omega_0\} \cup (\cup_{i \in I} BH_i(\omega_0, t^c)) = \Omega$$

which establishes that $\left(\Omega^c, \omega_0, s, (t_i^c)_{i \in I}\right) = \left(\Omega, \omega_0, s, (t_i)_{i \in I}\right)$. ■

**Proof.** **[Proposition 9]** From Proposition 6, one can deduce that

$$\left(\Omega^c, \omega_0, s, (t_i^c)_{i \in I}\right) = \left(\widetilde{\Omega}, \omega_0, s, \left(\tilde{t}_i\right)_{i \in I}\right)$$

and $\forall \omega \in \Omega^c$, $t_i^c(\omega) = t_i(\omega) \cap \left(\cap_{j \in I^c} t_j\left(\omega_0\right)\right)$
The result readily follows. ■

**Proof.** **[Proposition 10]** This is readily deduced from three observations:

- After a communication, if $i \in I^c$, then by definition, for all $\omega \in \Omega^c$, $t_i^c(\omega) = t_i^c(\omega_0)$.
- If we start from a situation where the MBS is such that for $i$, $\forall \omega \in \Omega$, $t_i(\omega) = t_i(\omega_0)$, then after a communication, it is also the case that $\omega \in \Omega^c$, $t_i^c(\omega) = t_i^c(\omega_0)$.
- Reducing MBS at each stage to minimal MBS if necessary, does not affect the two previous properties.

Thus if the sequential communication is exhaustive, we have that $\forall \omega \in \Omega_T^c$, $t_{i,T}^c(\omega) = t_{i,T}^c(\omega_0)$ which characterizes CSBS. ■

**Proof.** **[Proposition 11]** This reaily deduced from the following observations:

- When the MBS is totally correct the first revision rule can be applied even if the MBS is not minimal. It yields the same MBS as if it were applied on the minimal MBS to begin with.
- At each stage, the revised MBS is totally correct, and the general revision rule is equivalent to the first rule.
- Therefore, revision can be done without worrying minimality of the MBS.
- Thus, the MBS eventually reached corresponds to taking the intersection of the all the agents' announcements, an operation that does not depend on the order of these announcements.

■

**Proof.** **[Proposition 12]**
Let $R(\Omega)$ be the set of representations of $\Omega$, i.e., the set of MBS $(\Omega', \omega_0', s', (t_i')_{i \in I})$ such that there exists a mapping $\sigma$ from $\Omega$ to $\Omega'$ that satisfies the properties of Definition 11.

Let $\overline{\sigma}$ be a mapping from $\Omega$ to $\Omega$ that satisfies $\overline{\sigma}(\omega_1) = \overline{\sigma}(\omega_2)$ if and only if there exists an MBS $(\Omega', \omega_0', s', (t_i')_{i\in I})$ and a mapping $\sigma$ from $\Omega$ to $\Omega'$ that satisfies the properties of Definition 11 such that $\sigma(\omega_1) = \sigma(\omega_2)$. Let $\overline{\Omega} = \overline{\sigma}(\Omega)$ and $\overline{\omega}_0 = \overline{\sigma}(\omega_0')$.

Define $\overline{s} : \overline{\Omega} \to S$ by $\overline{s}(\overline{\omega}) = s(\omega_1)$ where $\omega_1 \in \Omega$ is such that $\overline{\sigma}(\omega_1) = \overline{\omega}$. This is well defined since if $\overline{\sigma}(\omega_1) = \overline{\sigma}(\omega_2)$ we know that there exists $\sigma$ such that $\sigma(\omega_1) = \sigma(\omega_2)$ which implies that $s(\omega_1) = s(\omega_2)$ since $\Omega'$ is a representation of $\Omega$ via $\sigma$.

Next, we show that if $\overline{\sigma}(\omega_1) = \overline{\sigma}(\omega_2)$ then $\overline{\sigma}(t_i(\omega_1)) = \overline{\sigma}(t_i(\omega_2))$. Since $\overline{\sigma}(\omega_1) = \overline{\sigma}(\omega_2)$, it must be the case that there exists $\sigma$ such that $\sigma(\omega_1) = \sigma(\omega_2)$. Then, $\sigma(t_i(\omega_1)) = \sigma(t_i(\omega_2))$. Now, let $\overline{\omega} \in \overline{\sigma}(t_i(\omega_1))$. There exists $\omega_3 \in t_i'(\omega_1)$ such that $\overline{\sigma}(\omega_3) = \overline{\omega}$. Since $\sigma(\omega_3) \in \sigma(t_i(\omega_1)) = \sigma(t_i(\omega_2))$, there exists $\omega_4 \in t_i(\omega_2)$ such that $\sigma(\omega_3) = \sigma(\omega_4)$. Hence, $\overline{\sigma}(\omega_3) = \overline{\sigma}(\omega_4) \in \overline{\sigma}(t_i(\omega_2))$ and therefore $\overline{\omega} \in \overline{\sigma}(t_i(\omega_2))$ proving that $\overline{\sigma}(t_i(\omega_1)) \subset \overline{\sigma}(t_i(\omega_2))$. Similarly, the reverse inclusion holds and hence $\overline{\sigma}(t_i(\omega_1)) = \overline{\sigma}(t_i(\omega_2))$.

Finally, define $\overline{t}_i : \overline{\Omega} \to 2^{\overline{\Omega}}$ by $\overline{t}_i(\overline{\omega}) = \overline{t}_i(\overline{\sigma}(\omega)) = \overline{\sigma}(t_i(\omega))$ where $\omega \in \Omega$ is such that $\overline{\sigma}(\omega) = \overline{\omega}$. This is well defined since we showed that if $\overline{\omega}$ has two antecedents $\omega_1$ and $\omega_2$, $\overline{\sigma}(t_i(\omega_1)) = \overline{\sigma}(t_i(\omega_2))$.

We first show that $\left(\overline{\Omega}, \overline{\omega}_0, \overline{s}, \left(\overline{t}_i\right)_{i\in I}\right)$ so defined is an MBS. The two conditions to check are condition *(iii)* and *(v)* of Definition 1. Check first condition *(iii)* and let $\overline{\omega}_2 \in \overline{t}_i(\overline{\omega}_1)$. There exist $\omega_1$ and $\omega_2$ such that $\overline{\sigma}(\omega_1) = \overline{\omega}_1$ and $\overline{\sigma}(\omega_2) = \overline{\omega}_2$ and $\omega_2 \in t_i(\omega_1)$. Hence $t_i(\omega_1) = t_i(\omega_2)$ and therefore $\overline{\sigma}(t_i(\omega_1)) = \overline{\sigma}(t_i(\omega_2))$, i.e., $\overline{t}_i(\overline{\omega}_1) = \overline{t}_i(\overline{\omega}_2)$.

We next check that condition $(v')$ holds (by Proposition 1, this is equivalent to check condition *(v)* of definition 1 directly). Let $\overline{\omega} \in \overline{\Omega}$. By construction, there exists $\omega \in \Omega$ such that $\overline{\sigma}(\omega) = \overline{\omega}$. Thus, there exists $r$ finite and a sequence $\{i_k\}_{k=1}^{k=r}$ with $i_k \in I$ for all $i$ such that $\omega \in t_{i_1}(t_{i_2}(...(t_{i_r}(\omega_0))))$. Hence,

$$\overline{\sigma}(\omega) \in \overline{\sigma}[t_{i_1}(t_{i_2}(...(t_{i_r}(\omega_0))))]$$

Recall that $\overline{\sigma}(t_i(\omega)) = \overline{t}_i(\overline{\sigma}(\omega))$. Hence,

$$\overline{\sigma}[t_{i_1}(t_{i_2}(...(t_{i_r}(\omega_0))))] = \overline{t}_{i_1}(\overline{\sigma}[t_{i_2}(...(t_{i_r}(\omega_0)))])$$

and, eventually,

$$\overline{\sigma}[t_{i_1}(t_{i_2}(...(t_{i_r}(\omega_0))))] = \overline{t}_{i_1}(\overline{t}_{i_2}(...(\overline{t}_{i_r}(\overline{\sigma}\omega_0)))) = \overline{t}_{i_1}(\overline{t}_{i_2}(...(\overline{t}_{i_r}(\overline{\omega}_0))))$$

proving condition $(v')$ of Proposition 1. Observe that $\left(\overline{\Omega}, \overline{\omega}_0, \overline{s}, \left(\overline{t}_i\right)_{i\in I}\right)$ is a representation of $(\Omega, \omega_0, s, (t_i)_{i\in I})$, since $\overline{\sigma}$ satisfies the conditions of Definition 11.

We next want to show that $\left(\overline{\Omega}, \overline{\omega}_0, \overline{s}, \left(\overline{t}_i\right)_{i\in I}\right)$ is minimal. Assume this is not the case and that there exists a representation $\left(\widetilde{\Omega}, \widetilde{\omega}_0, \widetilde{s}, \left(\widetilde{t}_i\right)_{i\in I}\right)$ of $\left(\overline{\Omega}, \overline{\omega}_0, \overline{s}, \left(\overline{t}_i\right)_{i\in I}\right)$ and a mapping $\widetilde{\sigma} : \widetilde{\Omega} \to \overline{\Omega}$ such that $\widetilde{\sigma}(\overline{\omega}_1) = \widetilde{\sigma}(\overline{\omega}_2)$ for some $\overline{\omega}_1, \overline{\omega}_2 \in \overline{\Omega}$, $\overline{\omega}_1 \neq \overline{\omega}_2$. Let $\omega_1$ and $\omega_2$ in $\Omega$ be such that $\overline{\omega}_1 = \overline{\sigma}(\omega_1)$ and $\overline{\omega}_2 = \overline{\sigma}(\omega_2)$. It is easy to show that

$\left(\widetilde{\Omega}, \widetilde{\omega}_0, \widetilde{s}, \left(\widetilde{t}_i\right)_{i\in I}\right)$ is also a representation of $(\Omega, \omega_0, s, (t_i)_{i\in I})$ via the mapping $\widetilde{\sigma} \circ \overline{\sigma}$. Hence, $\widetilde{\Omega} \in R(\Omega)$ and $\overline{\sigma}(\omega_1) = \overline{\sigma}(\omega_2)$, i.e., $\overline{\omega}_1 = \overline{\omega}_2$, a contradiction. ■

**Proof.** [**Proposition 13**] By Proposition 12, $\Omega$ and $\Omega'$ have a common minimal representation $\Omega$". Let $\sigma : \Omega \to \Omega$" and $\sigma' : \Omega' \to \Omega$" be the associated mappings. By definition, $\sigma$ and $\sigma'$ are onto. Assume $\sigma$ is not one-to-one, i.e., there exist $\omega_1, \omega_2 \in \Omega$, $\omega_1 \neq \omega_2$, such that $\sigma(\omega_1) = \sigma(\omega_2)$. This implies that $\Omega$ is not minimal, a contradiction. Hence $\sigma$ is one-to-one. A similar argument holds for $\sigma'$. Therefore, $(\sigma')^{-1} \circ \sigma$ is a well defined mapping from $\Omega$ to $\Omega'$ that is one-to-one and onto. Take $\phi = (\sigma')^{-1} \circ \sigma$. Conditions $(i)$ to $(iv)$ hold by construction. ■

**Proof.** [**Proposition 14**] Observe that for all $i \in I$, $t'_i(\sigma(\omega)) = \sigma\left(t_i(\omega)\right)$ by construction and $\omega \in t_i(\omega)$ by assumption. Hence, $\sigma(\omega) \in \sigma\left(t_i(\omega)\right)$ and therefore $\sigma(\omega) \in t'_i(\sigma(\omega))$ for all $i \in I$. ■

**Proof.** [**Proposition 15**] Let $\phi$ be defined as in Proposition 13. Since the orders are compatible, it is easy to check that $\widetilde{t}'_i \circ \phi(\omega) = \phi \circ \widetilde{t}_i(\omega)$ for all $\omega \in \Omega$. Hence, $((\Omega')^c, \omega'_0, s', ((t'_i)^c)_{i\in I})$ is a representation of $(\Omega^c, \omega_0, s, (t^c_i)_{i\in I})$. ■

## References


ALCHOURRON, C., P. GARDENFORS, AND D. MAKINSON (1985): "On the logic of theory change: partial meet contraction and revision functions," *The Journal of Symbolic Logic*, 50, 510–530.

AUMANN, R. (1976): "Agreeing to disagree," *Annals of Statistics*, 4(6), 1236–1239.

BACHARACH, M. (1985): "Some extensions of a claim of Aumann in an axiomatic model of knowledge," *Journal of Economic Theory*, 37, 161–190.

BONANNO, G., AND K. NEHRING (1998): "Assessing the truth axiom under incomplete information," *Mathematical Social Sciences*, 38, 3–29.

GEANAKOPLOS, J. (1994): "Common knowledge," in *Handbook of game theory*, ed. by R. Aumann, and S. Hart, vol. 2, chap. 40, pp. 1437–1496. North Holland - Amsterdam.

GEANAKOPLOS, J., AND H. POLEMARCHAKIS (1982): "We can't disagree forever," *Journal of Economic Theory*, 28, 192–200.

GUL, F., AND R. LUNDHOLM (1995): "Endogenous timing and the clustering of agents' decisions," *Journal of Political Economy*, 103, 1039–1066.

MAKINSON, D. (1985): "How to give it up: a survey of some formal aspects of the logic of theory change," *Synthese*, 62, 347–363.

MERTENS, J.-F., AND S. ZAMIR (1985): "Formulation of Bayesian analysis for games with incomplete information," *International Journal of Game Theory*, 10, 619–632.

MILGROM, P., AND N. STOKEY (1982): "Information, trade and common knowledge," *Journal of Economic Theory*, 26, 17–27.

MORRIS, S. (1994): "Trade with heterogeneous prior beliefs and asymmetric information," *Econometrica*, 62, 1327–1347.

RUBINSTEIN, A. (1989): "The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge," *Amercian Economic Review*, 79, 385–391.

SAMET, D. (1990): "Ignoring ignorance and agreeing to disagree," *Journal of Economic Theory*, 52, 190–207.