

Bonanno, Giacomo; Nehring, Klaus; Marcellino, Massimiliano

Working Paper

Varieties of Interpersonal Compatibility of Beliefs

Working Paper, No. 99-2

Provided in Cooperation with:

University of California Davis, Department of Economics

Suggested Citation: Bonanno, Giacomo; Nehring, Klaus; Marcellino, Massimiliano (1999) : Varieties of Interpersonal Compatibility of Beliefs, Working Paper, No. 99-2, University of California, Department of Economics, Davis, CA

This Version is available at:

<https://hdl.handle.net/10419/189486>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

VARIETIES OF INTERPERSONAL COMPATIBILITY OF BELIEFS

by

**Giacomo Bonanno
and
Klaus Nehring**

Working Paper Series No. 99-02
University of California, Davis
Department of Economics

March 1999



Department of Economics, University of California
One Shields Avenue, Davis, CA 95616-8578

Download from www.econ.ucdavis.edu/workingpapers/wpapers.html

VARIETIES OF INTERPERSONAL COMPATIBILITY OF BELIEFS

Giacomo Bonanno and Klaus Nehring

Department of Economics,
University of California,
Davis, CA 95616-8578 USA
gfbonanno@ucdavis.edu
kdnehring@ucdavis.edu

March 1999

Keywords: Common Belief, Negative Introspection,
Truth Axiom, Interpersonal Compatibility.

Abstract

The logic of common belief does not always reflect the logic of individual beliefs. In particular, the Negative Introspection property fails at the common belief level, that is, it can happen that neither is A commonly believed nor is it common belief that A is not commonly believed. Indeed Negative Introspection of common belief involves restrictions on individual beliefs of an intersubjective nature. We consider a number of interpersonal restrictions on beliefs and study their relationship. We also characterize Negative Introspection of common belief in terms of interpersonal properties of individual beliefs. All the results are proved syntactically.

1. Introduction

Since Lewis's (1969) and Aumann's (1976) pioneering contributions, the concepts of common knowledge and common belief have been discussed extensively in the literature, both syntactically and semantically¹. At the individual level the difference between knowledge and belief is usually identified with the presence or absence of the Truth Axiom ($\Box_i A \rightarrow A$), which is interpreted as "if individual i believes that A , then A is true". In such a case the individual is often said to *know* that A (thus it is possible for an individual to believe a false proposition but she cannot know a false proposition). Going to the interpersonal level, the literature then distinguishes between common knowledge and common belief on the basis of whether or not the Truth Axiom is postulated at the individual level. However, while at the individual level the Truth Axiom captures merely a relationship between the individuals' beliefs and the external world, at the interpersonal level it has very strong implications. For example, the following is a consequence of the Truth Axiom: $\Box_i \Box_j A \rightarrow \Box_i A$, that is, if individual i believes that individual j believes that A , then individual i herself believes that A . Thus, in contrast to other axioms, the Truth Axiom does not merely reflect individual agents' "logic of belief". (The reason why the Truth Axiom is much stronger in an interpersonal context than appears at first glance is that it amounts to assuming that agreement of any individual's belief with the truth is common knowledge). Given its logical force, it is not surprising to find that it has strong implications for the logic of common knowledge. In particular, if each individual's beliefs satisfy the strongest logic of knowledge (namely S5 or KT5), the associated common knowledge operator satisfies this logic too. Such is not the case for belief: bereft of the Truth Axiom, even the strongest logic for individual belief (KD45) is insufficient to ensure the satisfaction of the "Negative Introspection" axiom for common belief: $\neg \Box_* A \rightarrow \Box_* \neg \Box_* A$ (where \Box_* denotes the common belief operator). That is to say, it can happen that neither is A commonly believed nor is it common belief that A is not commonly believed. Indeed the Negative Introspection axiom for common belief implies restrictions on individual beliefs of an intersubjective nature. In this paper we consider a variety of intersubjective compatibility restrictions on the beliefs of the individuals and study their relationship. We also provide a characterization of Negative Introspection for common belief.

¹See, for example, Fagin *et al* (1995) and references therein.

2. The basic system \mathbf{K}_n^*

We consider a multimodal system with $n + 1$ operators $\Box_1, \Box_2, \dots, \Box_n, \Box_*$ where, for $i = 1, \dots, n$, the interpretation of $\Box_i A$ is “individual i believes that A ”, while $\Box_* A$ is interpreted as “it is common belief that A ”. The basic system \mathbf{K}_n^* is given by a suitable axiomatization of Propositional Calculus together with the following axiom schemata and rules of inference:

- K** $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$ $(\forall \Box \in \{\Box_1, \dots, \Box_n, \Box_*\})$
- CB1** $\Box_* A \rightarrow \Box_i A$ $(\forall i = 1, \dots, n)$
- CB2** $\Box_* A \rightarrow \Box_i \Box_* A$ $(\forall i = 1, \dots, n)$
- CB3** $\Box_*(A \rightarrow \Box_1 A \wedge \dots \wedge \Box_n A) \rightarrow (\Box_1 A \wedge \dots \wedge \Box_n A \rightarrow \Box_* A)$
- MP (modus ponens) $\frac{A, A \rightarrow B}{B}$
- RN (necessitation) $\frac{A}{\Box A}$ $(\forall \Box \in \{\Box_1, \dots, \Box_n, \Box_*\})$

We now turn to the semantics. A *standard model* is a tuple

$$\mathcal{M} = \langle W, R_1, \dots, R_n, R_*, V \rangle$$

where W is a non-empty set of *possible worlds*, R_1, \dots, R_n, R_* are binary *accessibility relations* on W and V is a *valuation*, that is, a function that associates with every atomic proposition p the set of possible worlds where p is true. The valuation is extended to the set of formulas in the usual way; we denote the fact that formula A is true at world w in model \mathcal{M} by $\mathcal{M}, w \models A$. Thus, in particular, for $i = 1, \dots, n$, $\mathcal{M}, w \models \Box_i A$ if and only if $\mathcal{M}, w' \models A$ for all w' such that $wR_i w'$. Similarly, $\mathcal{M}, w \models \Box_* A$ if and only if $\mathcal{M}, w' \models A$ for all w' such that $wR_* w'$. The following result is well-known (cf., for example, Bonanno, 1996).

Theorem 2.1. *The system \mathbf{K}_n^* is sound and complete with respect to the class of standard models where R_* is the transitive closure of $R_1 \cup \dots \cup R_n$.²*

A standard model where R_* is the transitive closure of $R_1 \cup \dots \cup R_n$ will be called a *CB-model*.

²That is, $aR_* b$ if and only if there are sequences $\langle w_1, \dots, w_m \rangle$ and $\langle i_1, \dots, i_m \rangle$ such that (1) $w_1 = a$, (2) $w_m = b$ and (3) for every $k = 1, \dots, m - 1$, $w_k R_{i_k} w_{k+1}$.

We will investigate extensions of \mathbf{K}_n^* obtained by adding one or more of the following axioms for individual beliefs:

- D** $\Box_i A \rightarrow \neg \Box_i \neg A$ ($i = 1, \dots, n$)
- T** $\Box_i A \rightarrow A$ ($i = 1, \dots, n$)
- 4** $\Box_i A \rightarrow \Box_i \Box_i A$ ($i = 1, \dots, n$)
- 4c** $\Box_i \Box_i A \rightarrow \Box_i A$ ($i = 1, \dots, n$)
- 5** $\neg \Box_i A \rightarrow \Box_i \neg \Box_i A$ ($i = 1, \dots, n$)

In the next section we list some theorems and derived rules of inference for \mathbf{K}_n^* which will be used later.

3. Some properties of \mathbf{K}_n^*

For every modal operator $\Box \in \{\Box_1, \dots, \Box_n, \Box_*\}$ we write \Diamond for $\neg \Box \neg$ (thus, for example, $\Diamond_i A$ stands for $\neg \Box_i \neg A$). Furthermore, PL stands for "Propositional Logic".

It is well-known (see, for example, Chellas, 1984, Lismont and Mongin, 1994) that \mathbf{K}_n^* has the following theorems and rules of inference:

- RK** $\frac{A \rightarrow B}{\Box A \rightarrow \Box B}$ for every $\Box \in \{\Box_1, \dots, \Box_n, \Box_*\}$
- RK \Diamond** $\frac{A \rightarrow B}{\Diamond A \rightarrow \Diamond B}$ for every $\Diamond \in \{\Diamond_1, \dots, \Diamond_n, \Diamond_*\}$
- M** $\Box(A \wedge B) \leftrightarrow \Box A \wedge \Box B$ for every $\Box \in \{\Box_1, \dots, \Box_n, \Box_*\}$
- R1CB** $\frac{A \rightarrow (\Box_1 A \wedge \dots \wedge \Box_n A)}{(\Box_1 A \wedge \dots \wedge \Box_n A) \rightarrow \Box_* A}$ (apply RN to the hypothesis, then use **CB3** and MP)
- R2CB** $\frac{A \rightarrow (\Box_1 A \wedge \dots \wedge \Box_n A)}{A \rightarrow \Box_* A}$ (apply R1CB to the hypothesis then use PL)

Proofs of the following lemma and corollary are given in the Appendix.

Lemma 1. *The following is provable in \mathbf{K}_n^* :*

$$\mathbf{CB4} \quad (\Box_1 A \wedge \dots \wedge \Box_n A) \wedge (\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A) \rightarrow \Box_* A$$

Corollary 1. *The following is provable in $\mathbf{K}_n^* + \mathbf{4c}$:*

$$\mathbf{CB5} \quad (\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A) \rightarrow \Box_* A$$

Remark 1. *It is well-known (cf. Chellas, 1984) that axiom **4c** is provable in $\mathbf{K}_n^* + \mathbf{5}$. Hence **CB5** is provable in $\mathbf{K}_n^* + \mathbf{5}$.*

4. Interpersonal compatibility of beliefs

In general, the common belief operator \Box_* does not inherit all the properties of the individuals' belief operators. Consider, for example, the counterpart for common belief of axiom **5** for individual beliefs:

$$\mathbf{5}^* \quad \neg\Box_*A \rightarrow \Box_*\neg\Box_*A.$$

Now, $\mathbf{5}^*$ is *not* provable in the system obtained by adding **D**, **4** and **5** to \mathbf{K}_n^* , which will be denoted by $\mathbf{K}_n^* + \mathbf{D45}$ (*a fortiori* it is not provable in a weaker system such as $\mathbf{K}_n^* + \mathbf{5}$). This is shown in the following example.

Example 1. Consider the following CB-model: $W = \{a, b\}$, $R_1 = \{(a, a), (b, b)\}$, $R_2 = \{(a, b), (b, b)\}$. Thus $R_* = \{(a, a), (a, b), (b, b)\}$. Let p be an atomic proposition which is true at b and false at a . Then the formula $(\neg\Box_*p \rightarrow \Box_*\neg\Box_*p)$, which is an instance of $\mathbf{5}^*$, is false at a . Since the system $\mathbf{K}_n^* + \mathbf{D45}$ is sound with respect to the class of CB-models where R_i is serial, transitive and euclidean ($i = 1, \dots, n$) and this model satisfies these properties, it follows that $\mathbf{5}^*$ is not provable in $\mathbf{K}_n^* + \mathbf{D45}$.

It follows from Example 1 that axiom $\mathbf{5}^*$ must involve further restrictions on the beliefs of the individuals which presumably are intersubjective in nature. In this section we consider various requirements of intersubjective compatibility of beliefs and study their relationship.

The following axioms capture interpersonal restrictions of various strength on the beliefs of the individuals.

$$\begin{array}{ll} \mathbf{C} & \Box_i A \rightarrow \Diamond_j A \\ \mathbf{TN} & \Box_i \Box_j A \rightarrow \Box_j A \\ \mathbf{TP} & \Box_i \Diamond_j A \rightarrow \Diamond_j A \\ \mathbf{IN} & \Box_i \Box_j A \rightarrow \Box_i A \\ \mathbf{IP} & \Diamond_i \Box_j A \rightarrow \Diamond_i A \\ \mathbf{SW} & \Box_i \Box_* A \rightarrow \Box_j \Box_* A \\ \mathbf{T}^{\text{CB}} & \Box_i \Box_* A \rightarrow \Box_* A \end{array}$$

C rules out the possibility that one individual believes A and at the same time some other individual believes $\neg A$. **TN** requires individuals to be correct in their beliefs about what others believe, while **TP** requires them to be correct in their beliefs about what others consider possible. By **IN** individuals must share the beliefs that they attribute to others, while **IP** is a weakening of this: if individual i considers it possible that individual j believes A then i himself must at least consider A possible. **SW** (for *Shared Worlds*) requires the individuals' beliefs about what is commonly believed to agree. Finally **T^{CB}** requires individuals to be correct in their beliefs about common belief.

Remark 2. *T*, the Truth Axiom for individuals beliefs, plays a much stronger role in multi-agent contexts than in single-agent ones. For a single individual the Truth Axiom captures merely a relationship between her beliefs and the external world; at the interpersonal level, on the other hand, it implies strong intersubjective compatibility of beliefs. Indeed, all of the above axioms (**C**, **TN**, **TP**, **IN**, **IP**, **SW** and **T^{CB}**) are provable in $\mathbf{K}_n^* + \mathbf{T}$. The reason why the Truth Axiom has strong implications in an interpersonal context is that it amounts to assuming that agreement of any individual's beliefs with the truth is commonly believed.

The following theorem establishes the relationship between the first five axioms, namely **C**, **TN**, **TP**, **IN** and **IP**.

Theorem 4.1.

- (i) $\mathbf{K}_n^* + \mathbf{IP} + \mathbf{TP} \vdash \mathbf{C}$
- (ii) $\mathbf{K}_n^* + \mathbf{4} + \mathbf{C} \vdash \mathbf{IP}$
- (iii) $\mathbf{K}_n^* + \mathbf{4} + \mathbf{C} \vdash \mathbf{TP}$
- (iv) $\mathbf{K}_n^* + \mathbf{5} + \mathbf{C} \vdash \mathbf{TN}$

Proof. The proof of (i) is as follows.

1. $\diamond_i \Box_j \neg A \rightarrow \neg \Box_i A$ (**IP**)
2. $\Box_i A \rightarrow \Box_i \diamond_j A$ (1,PL)
3. $\Box_i \diamond_j A \rightarrow \diamond_j A$ (**TP**)
4. $\Box_i A \rightarrow \diamond_j A$ (2,3,PL)

(ii) Next we prove that $\mathbf{K}_n^* + \mathbf{4} + \mathbf{C} \vdash \mathbf{IP}$:

1. $\Box_i \neg A \rightarrow \neg \Box_j A$ (**C**)
2. $\Box_i \Box_i \neg A \rightarrow \Box_i \neg \Box_j A$ (1,RK)
3. $\Box_i \neg A \rightarrow \Box_i \Box_i \neg A$ (axiom 4)
4. $\Box_i \neg A \rightarrow \Box_i \neg \Box_j A$ (2,3,PL)
5. $\diamond_i \Box_j A \rightarrow \diamond_i A$ (4,PL)

(iii) The proof of (iii) is as follows:

1. $\Box_j \neg A \rightarrow \Box_j \Box_j \neg A$ (axiom 4)
2. $\Box_j \Box_j \neg A \rightarrow \neg \Box_i \neg \Box_j \neg A$ (C)
3. $\Box_j \neg A \rightarrow \neg \Box_i \neg \Box_j \neg A$ (1,2,PL)
4. $\Box_i \Diamond_j A \rightarrow \Diamond_j A$ (3,PL)

(iv) The proof that $\mathbf{K}_n^* + \mathbf{5} + \mathbf{C} \vdash \mathbf{TN}$ is as follows:

1. $\neg \Box_j A \rightarrow \Box_j \neg \Box_j A$ (axiom 5)
2. $\Box_j \neg \Box_j A \rightarrow \neg \Box_i \Box_j A$ (C)
3. $\neg \Box_j A \rightarrow \neg \Box_i \Box_j A$ (1,2,PL)
4. $\Box_i \Box_j A \rightarrow \Box_j A$ (3,PL)

■

By Theorem 4.1, Compatibility (axiom C) has strong implications for the truth of intersubjective beliefs (TP and TN); moreover, it crucially involves some form of intersubjective agreement (IP). However, as the following example shows, intersubjective truth alone (TP and TN) fails to imply any intersubjective agreement (IP or IN) and therefore fails to imply C.

Example 2. Consider the following model: $W = \{a, b\}$, $R_1 = \{(a, b), (b, b)\}$, $R_2 = \{(a, a), (b, a)\}$. Thus R_* (the transitive closure of $R_1 \cup R_2$) is the universal relation. Let p be an atomic proposition which is true at a and false at b . Then at a all of the following are true $\Box_1 \Box_2 p$, $\Diamond_1 \Box_2 p$, $\neg \Box_1 p$, $\Box_1 \neg p$, $\Box_2 p$, thus falsifying IN, IP, and C. On the other hand, both TP and TN are true at every world.³

Similarly, the following example shows that intersubjective agreement (IP and IN) does not imply intersubjective truth (TP and TN) nor does it imply compatibility (C).

Example 3. Let $W = \{a, b, c\}$, $R_1 = \{(a, a), (b, a), (c, c)\}$, $R_2 = \{(a, a), (b, c), (c, c)\}$. Thus $R_* = \{(a, a), (b, a), (b, c), (c, c)\}$. Let p be an atomic proposition which is true at a and false at b and c . Here at world b all the following are true: $\Box_1 \Box_2 p$, $\Box_1 \Diamond_2 p$, $\neg \Box_2 p$, $\Box_2 \neg p$ and $\Box_1 p$. Thus at b TP, TN and C are falsified. On the other hand both IP and IN are valid in this model.⁴

³This is a consequence of the following fact, which is proved in the appendix. Axioms TN and TP are valid in the class of CB-models where, $\forall i, j \in N, \forall w, w', w'' \in W$, (1) R_i is serial and (2) if $wR_i w'$ and $w'R_j w''$ then $wR_j w''$. The model of the above example satisfies this property.

⁴This is a consequence of the following fact, which is proved in the Appendix. Axioms IN and IP are valid in the class of CB-models where, $\forall i, j \in N, \forall w \in W, \exists w' \in W$ such that (1) $wR_i w'$ and (2) $\forall w'' \in W$, if $wR_i w''$ then $w'R_j w''$. The model of the above example satisfies this property.

We now turn to the relationship between **SW** and \mathbf{T}^{CB} .

Theorem 4.2. (i) $\mathbf{K}_n^* + \mathbf{T}^{CB} \vdash \mathbf{SW}$
(ii) $\mathbf{K}_n^* + \mathbf{4c} + \mathbf{SW} \vdash \mathbf{T}^{CB}$

Proof. For (i)

1. $\Box_i \Box_* A \rightarrow \Box_* A$ (\mathbf{T}^{CB})
2. $\Box_* A \rightarrow \Box_j \Box_* A$ (**CB2**)
3. $\Box_i \Box_* A \rightarrow \Box_j \Box_* A$ (1,2,PL)

For (ii)

1. $\Box_i \Box_* A \rightarrow \Box_1 \Box_* A$ (**SW**)
- (**SW**)
- n. $\Box_i \Box_* A \rightarrow \Box_n \Box_* A$ (**SW**)
- n+1. $\Box_i \Box_* A \rightarrow (\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A)$ (1,...,n,PL)
- n+2. $(\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A) \rightarrow \Box_* A$ (**CB5**: cf. Corollary 1)
- n+3. $\Box_i \Box_* A \rightarrow \Box_* A$ (n+1, n+2, PL)

■

The next theorem relates \mathbf{T}^{CB} to $\mathbf{5}^*$.

Theorem 4.3. (i) $\mathbf{K}_n^* + \mathbf{D} + \mathbf{5}^* \vdash \mathbf{T}^{CB}$
(ii) $\mathbf{K}_n^* + \mathbf{5} + \mathbf{T}^{CB} \vdash \mathbf{5}^*$

Proof. The proof of (i) is as follows:

1. $\neg \Box_* A \rightarrow \Box_* \neg \Box_* A$ (axiom $\mathbf{5}^*$)
2. $\Box_* \neg \Box_* A \rightarrow \Box_i \neg \Box_* A$ (**CB1**)
3. $\Box_i \neg \Box_* A \rightarrow \neg \Box_i \Box_* A$ (**D**)
4. $\neg \Box_* A \rightarrow \neg \Box_i \Box_* A$ (1,2,3,PL)
5. $\Box_i \Box_* A \rightarrow \Box_* A$ (4,PL)

Next we prove (ii).

1. $\Box_i \Box_* A \rightarrow \Box_* A$ (\mathbf{T}^{CB})
2. $\neg \Box_* A \rightarrow \neg \Box_i \Box_* A$ (1,PL)
3. $\neg \Box_i \Box_* A \rightarrow \Box_i \neg \Box_i \Box_* A$ (axiom **5**)
4. $\neg \Box_* A \rightarrow \Box_i \neg \Box_i \Box_* A$ (2,3,PL)
5. $\Box_* A \rightarrow \Box_i \Box_* A$ (**CB2**)
6. $\neg \Box_i \Box_* A \rightarrow \neg \Box_* A$ (5,PL)
7. $\Box_i \neg \Box_i \Box_* A \rightarrow \Box_i \neg \Box_* A$ (6,RK)
8. $\neg \Box_* A \rightarrow \Box_i \neg \Box_* A$ (4,7,PL)

Now, a repetition of steps 1-8 for every $i = 1, \dots, n$ leads to (by 8 and PL)

9. $\neg\Box_*A \rightarrow (\Box_1\neg\Box_*A \wedge \dots \wedge \Box_n\neg\Box_*A)$
10. $\neg\Box_*A \rightarrow \Box_*\neg\Box_*A$ (9, R2CB)

■

Remark 3. Since (cf. Remark 1) $\mathbf{K}_n^* + \mathbf{5} \vdash \mathbf{4c}$, it follows from (ii) of Theorem 4.2 that $\mathbf{K}_n^* + \mathbf{5} + \mathbf{SW} \vdash \mathbf{T}^{CB}$. Thus, by Theorem 4.3, $\mathbf{K}_n^* + \mathbf{5} + \mathbf{SW} \vdash \mathbf{5}^*$. Furthermore, by Theorems 4.2 and 4.3, $\mathbf{K}_n^* + \mathbf{D} + \mathbf{5}^* \vdash \mathbf{SW}$. It follows from Theorem 4.3 that $\mathbf{K}_n^* + \mathbf{D5} + \mathbf{SW}$ and $\mathbf{K}_n^* + \mathbf{D5} + \mathbf{5}^*$ are the same system.

We now turn to the relationship between axioms **TN**, **TP**, **IN**, **IP**, **C** and the Shared Worlds axiom (**SW**).

Proposition 1. None of **TN**, **TP**, **IN**, **IP** and **C** is provable in $\mathbf{K}_n^* + \mathbf{D45} + \mathbf{SW}$ (thus, a fortiori, in a weaker system such as $\mathbf{K}_n^* + \mathbf{SW}$).

Proof. First of all, it is straightforward to show that the system $\mathbf{K}_n^* + \mathbf{D45} + \mathbf{SW}$ is sound with respect to the class of CB-models where (1) $\forall i \in N$, R_i is serial, transitive and euclidean, and (2) $\forall i, j \in N$, $\forall x, y, z \in W$, if xR_jy and yR_*z then $\exists w \in W$ such that xR_iw and wR_*z . The following model belongs to this class: $W = \{a, b, c\}$, $R_1 = \{(a, b), (b, b), (c, c)\}$, $R_2 = \{(a, a), (b, c), (c, c)\}$, $R_3 = \{(a, a), (b, b), (c, a)\}$; thus R_* is the universal relation. Let p be an atomic proposition which is true at c and false at a and b . Then this model validates all the theorems of $\mathbf{K}_n^* + \mathbf{D45} + \mathbf{SW}$ (in particular **SW** itself)⁵. It is easily checked that

- $\Box_1\Box_2p \rightarrow \Box_2p$ (which is an instance of **TN**) is *false* at a
- $\Box_1\Diamond_2p \rightarrow \Diamond_2p$ (which is an instance of **TP**) is *false* at a
- $\Box_1\Box_2p \rightarrow \Box_1p$ (which is an instance of **IN**) is *false* at a
- $\Diamond_1\Box_2p \rightarrow \Diamond_1p$ (which is an instance of **IP**) is *false* at a
- $\Box_2p \rightarrow \neg\Box_3\neg p$ (which is an instance of **C**) is *false* at b

■

Theorem 4.4. (i) $\mathbf{K}_n^* + \mathbf{TN} \vdash \mathbf{SW}$
(ii) $\mathbf{K}_n^* + \mathbf{D5} + \mathbf{TP} \vdash \mathbf{SW}$

⁵Since R_* is the universal relation, for every formula A and every world w , $w \models \Box_*A$ if and only if A is valid (i.e. true at every world). It follows from seriality of R_i that if individual i believed that A is common belief then A is indeed commonly believed (thus **T^{CB}** is valid). Hence everybody shares i 's belief that A is common belief (thus **SW** is valid).

Proof. For (i)

1. $\Box_* A \rightarrow \Box_j \Box_* A$ (CB2)
2. $\Box_i \Box_* A \rightarrow \Box_i \Box_j \Box_* A$ (1,RK)
3. $\Box_i \Box_j \Box_* A \rightarrow \Box_j \Box_* A$ (TN)
4. $\Box_i \Box_* A \rightarrow \Box_j \Box_* A$ (2,3,PL)

For (ii)

1. $\Box_* A \rightarrow \Box_j \Box_* A$ (CB2)
2. $\Diamond_j \Box_* A \rightarrow \Diamond_j \Box_j \Box_* A$ (1,RK \Diamond)
3. $\neg \Box_j \Box_* A \rightarrow \Box_j \neg \Box_j \Box_* A$ (axiom 5)
4. $\Diamond_j \Box_j \Box_* A \rightarrow \Box_j \Box_* A$ (3,PL)
5. $\Diamond_j \Box_* A \rightarrow \Box_j \Box_* A$ (2,4,PL)
6. $\Box_j \Box_* A \rightarrow \Diamond_j \Box_* A$ (axiom D)
7. $\Box_i \Box_* A \rightarrow \Box_i \Box_j \Box_* A$ (1,RK)
8. $\Box_i \Box_j \Box_* A \rightarrow \Box_i \Diamond_j \Box_* A$ (6,RK)
9. $\Box_i \Box_* A \rightarrow \Box_i \Diamond_j \Box_* A$ (7,8,PL)
10. $\Box_i \Diamond_j \Box_* A \rightarrow \Diamond_j \Box_* A$ (TP)
11. $\Box_i \Box_* A \rightarrow \Diamond_j \Box_* A$ (9,10,PL)
12. $\Box_i \Box_* A \rightarrow \Box_j \Box_* A$ (5,11,PL)

■

5. Concluding remarks

We have considered a variety of interpersonal compatibility restrictions on the beliefs of the individuals. The strongest restrictions are obtained by imposing the Truth Axiom on individual beliefs because of the implied common knowledge of the correctness of beliefs. Of the remaining conditions **C** is the strongest, since it implies all the others. The weakest of all is **SW**, which in the system $\mathbf{K}_n^* + \mathbf{D5}$ turns out to be equivalent to **5*** Negative Introspection of Common Belief. The relationship among all the axioms considered in this paper (**T**, **C**, **TN**, **TP**, **IN**, **IP**, **SW**, **T^{CB}**, **5***) is summarized in Figure 1.

Weaker intersubjective restrictions on beliefs can be obtained from **C**, **TN**, **TP**, **IN**, **IP** and **SW** by replacing \Box_j with \Box_* . In this case **TN** becomes **T^{CB}**, while the remaining axioms become

$$\begin{array}{ll}
\mathbf{C}^* & \Box_i A \rightarrow \Diamond_* A \\
\mathbf{TP}^* & \Box_i \Diamond_* A \rightarrow \Diamond_* A \\
\mathbf{IN}^* & \Box_i \Box_* A \rightarrow \Box_i A \\
\mathbf{IP}^* & \Diamond_i \Box_* A \rightarrow \Diamond_i A \\
\mathbf{SW}^* & \Box_i \Box_* A \rightarrow \Box_* \Box_* A
\end{array}$$

It is straightforward to show that all of the above are provable in $\mathbf{K}_n^* + \mathbf{D45}$.⁶

⁶Indeed, \mathbf{C}^* and \mathbf{TP}^* are provable in $\mathbf{K}_n^* + \mathbf{D}$, \mathbf{IN}^* is provable in $\mathbf{K}_n^* + \mathbf{4c}$ and \mathbf{IP}^* is provable in $\mathbf{K}_n^* + \mathbf{D4}$ or in $\mathbf{K}_n^* + \mathbf{D5}$.

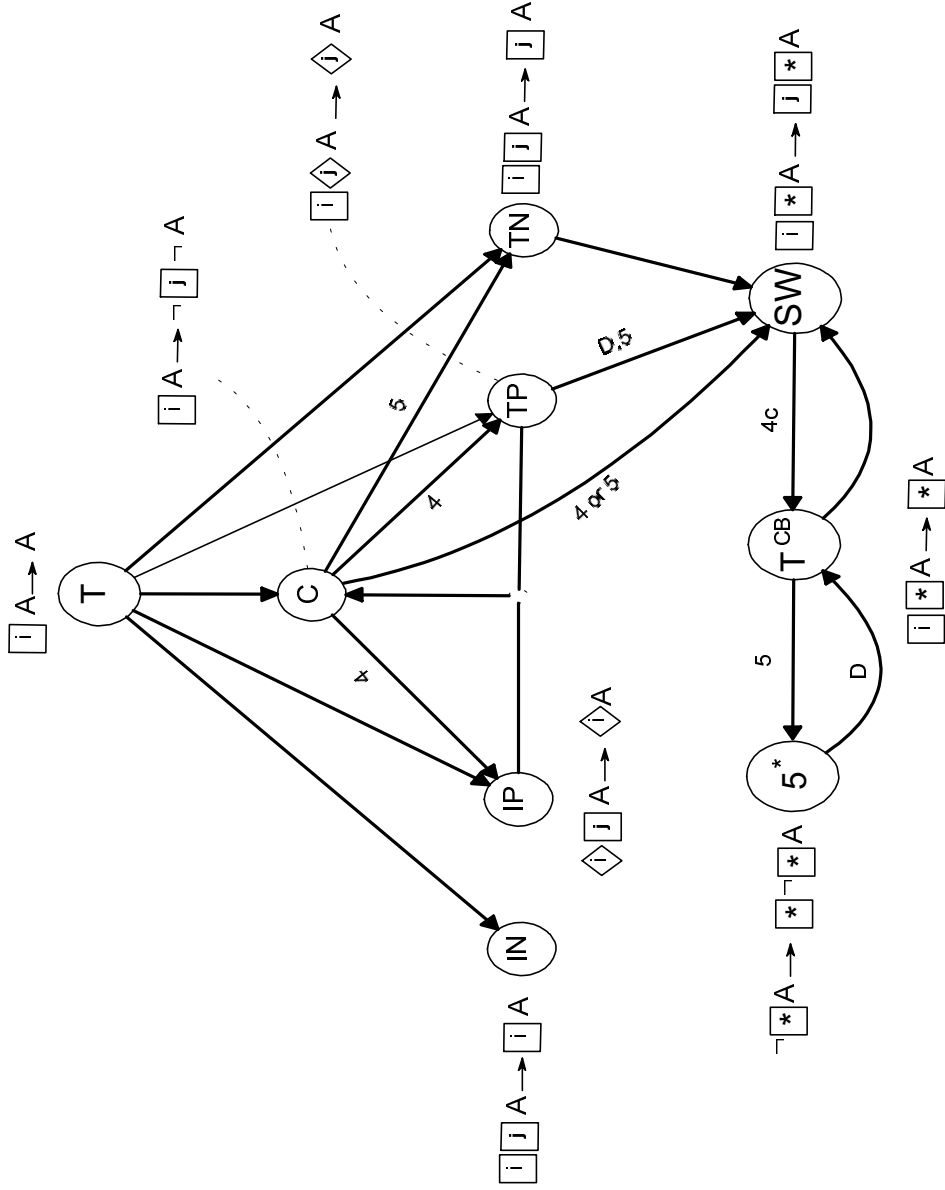


Figure 1

6. Appendix

Proof of Lemma 1.

1. $(\Box_1 A \wedge \dots \wedge \Box_n A) \wedge (\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A) \rightarrow \Box_1(A \wedge \Box_* A) \wedge \dots \wedge \Box_n(A \wedge \Box_* A)$ (M,PL)
2. $\Box_* A \rightarrow \Box_1 A \wedge \dots \wedge \Box_n A$ (CB1,PL)
3. $\Box_* A \rightarrow \Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A$ (CB2,PL)
4. $\Box_* A \rightarrow \Box_1(A \wedge \Box_* A) \wedge \dots \wedge \Box_n(A \wedge \Box_* A)$ (1,2,3,PL)
5. $A \wedge \Box_* A \rightarrow \Box_* A$ (PL)
6. $A \wedge \Box_* A \rightarrow \Box_1(A \wedge \Box_* A) \wedge \dots \wedge \Box_n(A \wedge \Box_* A)$ (4,5,PL)
7. $\Box_1(A \wedge \Box_* A) \wedge \dots \wedge \Box_n(A \wedge \Box_* A) \rightarrow \Box_*(A \wedge \Box_* A)$ (6, R1CB)
8. $\Box_*(A \wedge \Box_* A) \rightarrow \Box_* A \wedge \Box_* \Box_* A$ (M)
9. $(\Box_1 A \wedge \dots \wedge \Box_n A) \wedge (\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A) \rightarrow \Box_* A \wedge \Box_* \Box_* A$ (1,7,8,PL)
10. $\Box_* A \wedge \Box_* \Box_* A \rightarrow \Box_* A$ (PL)
11. $(\Box_1 A \wedge \dots \wedge \Box_n A) \wedge (\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A) \rightarrow \Box_* A$ (9,10,PL)

■

Proof of Corollary 1.

1. $\Box_* A \rightarrow \Box_i A$ (CB1)
2. $\Box_i \Box_* A \rightarrow \Box_i \Box_i A$ (1,RK)
3. $\Box_i \Box_i A \rightarrow \Box_i A$ (axiom 4c)
4. $\Box_i \Box_* A \rightarrow \Box_i A$ (2,3,PL)
5. $(\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A) \rightarrow (\Box_1 A \wedge \dots \wedge \Box_n A)$ (4,PL)
6. $(\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A) \rightarrow (\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A)$ (PL)
7. $(\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A) \rightarrow (\Box_1 A \wedge \dots \wedge \Box_n A) \wedge (\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A)$ (5,6,PL)
8. $(\Box_1 A \wedge \dots \wedge \Box_n A) \wedge (\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A) \rightarrow \Box_* A$ (CB4)
9. $(\Box_1 \Box_* A \wedge \dots \wedge \Box_n \Box_* A) \rightarrow \Box_* A$ (7,8,PL)

■

Lemma 2. *Axioms TN and TP are valid in the class of CB-models where, $\forall i, j \in N, \forall w, w', w'' \in W$, (1) R_i is serial and (2) if $wR_i w'$ and $w'R_j w''$ then $wR_j w''$.*

Proof. First we prove validity for **TN**. Fix an arbitrary model that satisfies the above two properties. Fix arbitrary i, j, a and an arbitrary formula A . Suppose that $a \models \Box_i \Box_j A$. By seriality of R_i there exists a b such that $aR_i b$. Fix an arbitrary such b . Then $b \models \Box_j A$. By seriality of R_j , there exists a c such that $bR_j c$. Fix an arbitrary such c . Then $c \models A$. By property (2) $aR_j c$. Hence $a \models \Box_j A$.

(2) Next we prove validity of **TP**. Fix an arbitrary model that satisfies the above two properties. Fix arbitrary i, j, a and an arbitrary formula A . Suppose that $a \models \Box_i \Diamond_j A$. By seriality of R_i there exists a b such that $aR_i b$. Then $b \models \Diamond_j A$. Hence there exists a c such that $bR_j c$ and $c \models A$. By property (2) $aR_j c$. Hence $a \models \Diamond_j A$. ■

Lemma 3. *Axioms **IN** and **IP** are valid in the class of CB-models where, $\forall i, j \in N, \forall w \in W, \exists w' \in W$ such that (1) $wR_i w'$ and (2) $\forall w'' \in W$, if $wR_i w''$ then $w'R_j w''$.*

Proof. First we prove validity for **IN**. Fix an arbitrary model that satisfies the above properties. Fix arbitrary i, j, a and an arbitrary formula A . Suppose that $a \models \Box_i \Box_j A$. By the assumed properties, there exists a b such that $aR_i b$. Choose an arbitrary such b . Then $b \models \Box_j A$. Choose an arbitrary c such that $aR_j c$. By the assumed properties, $bR_j c$. Hence $c \models A$. Therefore $a \models \Box_i A$.

(2) Next we prove validity for **IP**. Fix an arbitrary model that satisfies the above properties. Fix arbitrary i, j, a and an arbitrary formula A . Suppose that $a \models \Diamond_i \Box_j A$. Then there exists a b such that $aR_i b$ and $b \models \Box_j A$. By the assumed properties (choosing $c = b$), $bR_j b$. Hence $b \models A$. Therefore, $a \models \Diamond_i A$. ■

References

- [1] Aumann, Robert. 1976. Agreeing to disagree, *Annals of Statistics*, 4, 1236-1239.
- [2] Bonanno, Giacomo. 1996. ‘On the logic of common belief’. *Mathematical Logic Quarterly*, 42: 305-311.
- [3] Chellas, Brian. 1984. *Modal logic: an introduction*. Cambridge University Press.
- [4] Fagin, Ronald, Joseph Halpern, Yoram Moses and Moshe Vardi. 1995, *Reasoning about knowledge*. MIT Press.
- [5] Lewis, David. 1969. *Convention: a philosophical study*, Harvard University Press.
- [6] Lismont, Luc and Philippe Mongin. 1994. ‘On the logic of common belief and common knowledge’, *Theory and Decision*, 37: 75-106.