Common Belief and Common Knowledge

Georg Meggle

1. Convictions are always convictions held by a certain person. However, we can also speak of common convictions held by different people, and of one and the same conviction being shared by different people. But then we are actually talking about the *proposition* each of these persons *believes* to be the respective content of their convictions, not (only) the *proposition of being convinced* experienced by the persons concerned. In other words, what we are talking about is the proposition A which is believed to be true by both X and Y, and not the propositions that X believes A or that Y believes A, respectively (in symbols: $B(X, A)$ and $B(Y, A)$). The same goes for common knowledge. All the fact that X and Y have certain common convictions and common knowledge means is that for certain propositions A it holds that A is the content of the (correct) convictions expressed by $B(X, A)$ and $B(Y, A)$.

We obtain more narrow concepts of *Common Belief and Knowledge* that A if we do not just require (1) everyone (in the relevant reference group or population P) to believe or know that A, but also (2) everyone (in P) to be aware that (1), and (3) everyone to be aware of (2), etc.

Such interpersonal concepts of belief and knowledge play an important part in each and every examination of *Social Facts*. There is a very simple reason for this: Social Facts (with respect to P) only become such when they are Commonly held to be true (in P). If we express a Common Belief within P that A by $CB(P, A)$, we can also say:

A is a possible Social Fact with respect to $P := A_P \leftrightarrow CB(P, A)$

To use more informal language, Social Facts are those with respect to which a consensus theory of truth would be correct.

Despite their relevance, concepts of Common Belief and Knowledge have only been deemed worthy of attention for a relatively short time. Even in the more recent literature, the logical structure of these concepts remains largely unexplained. The following proposals are intended to make up for this shortcoming.[1]

---

[1] *These proposals have been developed in connection with Meggle (1984), see the appendix thereof. This appendix also contains the proofs for all the interpersonal belief theorems formulated herein. Concerning the symbols used: $\neg$, $\wedge$, $\vee$, $\supset$, $\equiv$ and Λ stand for negation, conjunction, adjunction, implication, equivalence and all-quantification; $\vdash$ is used for provability or derivability. For analytical conclusions, equivalencies,*

incompatibilities and contradictions we use: $\rightarrow$, $\leftrightarrow$, $\succ\!\!-\!\!\prec$, and $\gg\!\!-\!\!\ll$. Also note that, in order to highlight the lack of directly corresponding colloquial terms for the phenomena (Social Facts) of Common and Mutual Belief and Knowledge explicated here, the artificial terms for such Social Facts have been capitalised. Previous versions of this paper were presented at the XVI. DeutscherKongreß für Philosophie at Berlin, 1993, and at the Workshop "Collective Intentionality", Munich, June 1999.

2. My proposals admittedly remain within very narrow limits. The concept of belief used is the strongest possible – in two respects. In the following $B(X, A)$ stands for the so-called strong belief, i.e. for firm convictions in contrast to mere suppositions. And belief expressed in this way is a strongly rational belief. It is therefore governed by the following laws:

RB: $A \vdash B(X, A)$
B1: $B(X, A \supset \mathrm{B}) \supset (B(X, A) \supset B(X, \mathrm{B}))$
B2: $B(X, A) \supset \neg B(X, \neg A)$
B3: $\Lambda x(B(X, F(x))) \supset B(X, \Lambda x F(x))$
B4: $B(X, A) \supset B(X, B(X, A))$
B5: $\neg B(X, A) \supset B(X, \neg B(X, A))$

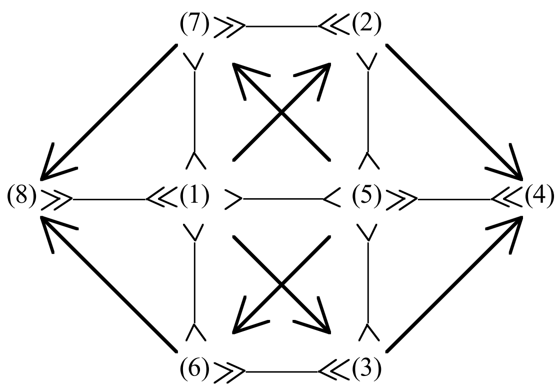
If we put $K(X, A) := B(X, A) \wedge A$ (in other words: X knows that A iff X is right in his conviction that A), the precisely corresponding basic laws also apply to knowledge – with B5 being the only exception. So much for the basics. (For more on this, cf. Kutschera (1976) and Lenzen (1980).)

3. Interpersonal states of belief are the cement which hold together our entire social relations. Nevertheless, as soon as we try to explain such states, our intuitive understanding of colloquial language runs into trouble unbelievably quickly – right at the very first fence.

Try it for yourself – if you like at the simplest of all possible interpersonality levels involving (i) just two people and (ii) merely whether one believes that the other believes something – or not. Compare the following:

(1)	$B(X, B(Y, A))$	(5)	$B(X, B(Y, \neg A))$
(2)	$B(X, \neg B(Y, \neg A))$	(6)	$B(X, \neg B(Y, A))$
(3)	$\neg B(X, \neg B(Y, A))$	(7)	$\neg B(X, \neg B(Y, \neg A))$
(4)	$\neg B(X, B(Y, \neg A))$	(8)	$\neg B(X, B(Y, A))$

Can you make head or tail of it? Well, through logical spectacles the states of belief ought to look like the diagram below:



4. Common Beliefs, even those shared by just two people, involve not only what one thinks of the other, but also what they both think about each other. For the 2-person case we could thus (e.g. with Schiffer (1989)) introduce a Common Belief as follows:

1st level: $B(X, A) \wedge B(Y, A)$
2nd level: $(\text{1st level}) \wedge B(X, B(Y, A)) \wedge B(Y, B(X, A))$
3rd level: $(\text{2nd level}) \wedge B(X, B(Y, B(X, A))) \wedge B(Y, B(X, B(Y, A)))$

etc.

5. More generally, i.e. for any size population P, where we write X belongs to P (or X is a member of P) as $X \in P$:

D1.a: $CB_1(P, A) := \Lambda X(X \in P \supset B(X, A))$
It is a Common Belief of the 1st level within P that A iff every member of P believes that A

D1.b: $CB_{n+1}(P, A) := CB_1(P, CB_n(P, A))$
It is a Common Belief of the n+1-th level within P that A iff it is a Common Belief of the 1st level within P that it is a Common Belief of the nth level within P that A

D1.c: $CB(P, A) := \Lambda n CB_n(P, A)$
It is a Common Belief within P that A iff within P it is a Common Belief at all levels that A

Accordingly, a Common Knowledge can be determined as follows:

D2.a: $CK_1(P, A) := \Lambda X(X \in P \supset K(X, A))$
D2.b: $CK_{n+1}(P, A) := CK_1(P, CK_n(P, A))$
D2.c: $CK(P, A) := \Lambda n CK_n(P, A)$

In direct analogy to the explanation of knowledge above (in §2) as a correct conviction, it therefore holds that:

(CK) $CK(P, A) \leftrightarrow CB(P, A) \wedge A$

It is a Common Knowledge within P that A iff it is a Common Belief within P that A, and A really is the case.

6. Between these concepts of Common Belief and Knowledge on the one hand and the basic concept of simple belief (always referring to a single person) on the other, there exists a whole range of interesting parallels which greatly simplify operating with these interpersonal concepts of belief.

The most important thing is that both Common Belief and Common Knowledge are governed by precisely the analogous laws as the concept of belief itself – up to the analogy to B5. For CB therefore (ditto then also for CK instead of CB):

RCB: $A \vdash CB(P, A)$
CB1: $CB(P, A \supset \mathrm{B}) \supset (CB(P, A) \supset CB(P, \mathrm{B}))$
CB2: $CB(P, A) \supset \neg CB(P, \neg A)$
CB3: $\Lambda x(CB(P, F(x))) \supset CB(P, \Lambda x F(x))$
CB4: $CB(P, A) \supset CB(P, CB(X, A))$

In particular it therefore holds that: If a B Theorem can be proved merely by using laws B1 to B4 and rule RB, then the analogous CB and CK sentences can also proved accordingly.

Hence, for instance, parallel to the simple Belief Theorems:

T.1: $A \supset B \vdash B(X, A) \supset B(X, B)$
T.2: $B(X, A) \wedge B(X, B) \supset B(X, A \wedge B)$
T.3: $B(X, A \equiv B) \supset (B(X, A) \equiv B(X, B))$

the following Cb and CK Theorems also apply:

T.B1: $A \supset B \vdash CB(P, A) \supset CB(P, B)$
T.B1*: $A \supset B \vdash CK(P, A) \supset CK(P, B)$
T.2: $CB(P, A) \wedge CB(P, B) \supset CB(P, A \wedge B)$
T.2*: $CK(P, A) \wedge CK(P, B) \supset CK(P, A \wedge B)$
T.3: $CB(P, A \equiv B) \supset (CB(P, A) \equiv CB(P, B))$
T.3*: $CK(P, A \equiv B) \supset (CK(P, A) \equiv CK(P, B))$

And similar to the connection between simple Belief and Knowledge, the following principles (among others) apply as well:

T.4: $B(X, A) \supset K(X, B(X, A))$
T.5: $B(X, A) \supset B(X, K(X, A))$
T.6: $K(X, A) \supset B(X,K(X, A))$

and also:

T.B4: $CB(P, A) \supset CK(P, CB(X, A))$
T.B5: $CB(P, A) \supset CB(P, CK(X, A))$
T.B6: $CK(P, A) \supset CB(P, CK(X, A))$

By contrast the parallel to the law (which assumes law B5):

T.7: $\neg B(X, A) \supset K(X, \neg B(X, A))$

i.e. therefore

(*) $\neg CB(P, A) \supset CK(P, \neg CB(P, A))$

does not hold. ($\neg CB(P, A)$ does indeed apply in the case of $P = \{X, Y\}$ for example if $\neg B(X, A)$; however, $B(Y, \neg B(X, A))$ by no means follows from this; and a fortiori neither, therefore, does $CK(P, \neg CB(P, A))$.)

7. By the way, directly corresponding to the simple law:

T.0: $B(X, A) \supset B(X, B(X, A))$

the following also holds:

T.G0: $CB(P,A) \supset CB(P, CB(P, A))$

A Common Belief in P is thus also a perfect example of a (in 1 above) so-called Social Fact w.r.t. P. (Correspondingly, a simple belief of X makes a perfect example of a – as can be analogously defined – strictly *Subjective Fact* w.r.t. the subject X . A is w.r.t. X *strictly subjective* := $A_X \leftrightarrow B(X,A)$.)

8. Concerning the Common Belief, too, we must also draw a distinction (as with simple belief) in the sense of specialisation of the customary *de re* vs. *de dicto* differentiation between a *generality in sensu composito* and a *generality in sensu diviso*. (i) corresponds to a former Common Belief; (ii) to a latter one:

(i) $CB(P, \Lambda x(B(x) \supset F(x)))$
It is a Common Belief within P that all B things are also F things.

(ii) $\Lambda x(B(x) \supset CB(P,F(x)))$
Of all B things, Common Belief prevails within P that they are also F things.

To be more precise: (i) and (ii) express a Common Belief which with respect to the B characteristic is generally *in sensu composito* and *in sensu diviso*.

(i) and (ii) would be equivalent assuming that the following also held:

$\Lambda x(B(x) \supset CB(P, B(x))) \wedge \Lambda x(\neg B) \supset CB(P, \neg B(x)))$, i.e.:
$\Lambda x(CK(P, B(x)) \vee CK(P, \neg B(x)))$ There is Common Belief in P concerning which things are B things and which are not.

As far as the (characteristic of) belonging to the group P is concerned, each of the CB_n concepts so far introduced is generally *in sensu composito* – as the 2nd level already showed. For $CB_2(P, A)$ means the same as:

(a) $\Lambda X(X \in P \supset B(X, \Lambda Y(Y \in P \supset B(Y, A))))$, i.e.
$CB_1(P, \Lambda Y(Y \in P \supset B(Y, A)))$
Everyone in P believes that everyone in P believes A

whereas a general Common Belief with respect to P-membership *in sensu diviso* would have to be expressed as:

(b) $\Lambda X \Lambda Y(X \in P \wedge Y \in P \supset B(X, B(Y, A)))$, i.e.:
$\Lambda Y(Y \in P \supset CB_1(P, B(Y, A)))$
Everyone in P believes of everyone in P that he believes A

(a) and (b) would be equivalent if

(c) $\Lambda X(CK_1(P, X \in P) \vee CK_1(P, X \notin P)$
Everyone from P knows who belongs to P and who doesn't.

Analogously, it holds that a general Common Belief with respect to P-membership *in sensu composito* is equivalent to the corresponding general Common Belief *in sensu diviso* if:

(c*) $\Lambda X(CK_1(P, X \in P) \vee CK_1(P, X \notin P)$
Common Knowledge prevails in P concerning who belongs to P and who doesn't.

Let P' be a subgroup of the population P. When then is a Common Belief in P also such in P'? The answer is: If Common Knowledge also prevails in P that P' is a subgroup of P.

9. Our doing and not doing often depend on what according to our conviction the others are doing; and we also know that the very same applies to the others as well. They too are often guided by what according to their convictions we are doing and not doing. However, it is also the case that we also know that the others convictions are sometimes incorrect – in which case, although we know their convictions, we do not share them. Therefore we cannot speak of Common Belief (as defined above). Such an interpersonal belief (which is weaker than a Common Belief) is involved in the following explicated *Mutual Belief*. (Although interpersonal concepts of belief and knowledge in the sense of D2 are found in the relevant literature, as far as I am aware they have not yet been defined there. Note that 'mutual knowledge' and 'common knowledge' are simply used by most authors as interchangeable terms for a Common Belief.)

D2.a: $MB_1(X, P, F(\hat{Y})) := B(X, \Lambda Y\,(Y \neq X \wedge Y \in P \supset F(Y)))$
From X's viewpoint, mutual belief of the 1st level exists in P that within P the characteristic $F(\hat{Y})$ exists iff X believes that everyone else in P has the property F.

D2.b: $MB_{n+1}(X, P, F(\hat{Y})) := MB_1(X, P, MB_n(\hat{Y}, P, F(\hat{Z})))$

D2.c: $MB(X, P, F(\hat{Y})) := \Lambda n MB_n(X, P, F(\hat{Y}))$

D2:d: $MB(P, F(\hat{Y})) := \Lambda X(X \in P \supset MB((X, P, F(\hat{Y})))$
There prevails mutual belief within P that the characteristic F exists in P.

Starting from these concepts, corresponding concepts of *Mutual Knowledge* can then be determined as the correct Mutual Convictions in each case.

In turn, there arise interesting parallels to the principles of simple belief:

RMB: $A \vdash MB(P, B(\hat{Y}, A))$
MB1: $MB(P, (F(\hat{Y}) \supset F^*(\hat{Y}))) \supset (MB(P, F(\hat{Y})) \supset MB(P, F^*(\hat{Y})))$
MB2: $MB(P, (F(\hat{Y}) \supset \neg MB(P, \neg F(\hat{Y}))$
MB3: $MB(P, (F(\hat{Y})) \supset MB(P, MB(\hat{X}, P, (F(\hat{Y})))$

And just like for Common Belief, the corresponding principles parallel to the simple belief theorems T1–T3 from Section 6 above also apply in turn to Mutual Belief.

10. A Mutual Belief is something weaker than a Common Belief. Therefore, my closing question is: How (and under what conditions) do we get from a Mutual Belief to a Common Belief and vice versa? In other words, how are Mutual and Common Belief related? The answer is, like this:

T.B7: $MB(P, F(\hat{Y})) \wedge MB(P, B(\hat{Y}, F(Y))) \wedge \Lambda X(X \in P \supset B(X, F(X)))$
$\equiv CB(P, \Lambda X(X \in P \supset F(X)))$

***

Bach, K., “Analytic Social Philosophy – Basic Concepts”, in J.Theory Soc.Behaviour 5, 1975, 182-214.
Gärdenfors, P., *Knowledge in Flux. Modeling the Dynamics of Epistemic States*, Cambridge, M.A., 1988.
Gilbert, M., *On Social Facts*, London, 1980.
Lewis, D., *Convention: A Philosophical Study*, Cambridge/Mass., 1969.
Kutschera, F. von, *Einführung in die intensionale Semantik*, Berlin, 1976.
Lenzen, W., *Glauben, Wissen und Wahrscheinlichkeit*, Wien / New York, 1980.
Meggle, G., *Handlungstheoretische Semantik*, 1984 (Habilitation-Ms.); to be published: Berlin / New York, 2000.
Schiffer, S., *Meaning*, Oxford, 1972.
Tuomela, R., *The Importance of Us: A Philosophical Study of Basic Social Notions*, Standford, 1995.