ON THE FACTIVITY OF IMPLICIT INTERSUBJECTIVE KNOWLEDGE

ALESSANDRO GIORDANI
Department of Philosophy, Catholic University of Milan

Abstract. The concept of knowledge can be modelled in epistemic modal logic and, if modelled by using a standard modal operator, it is subject to the problem of logical omniscience. The classical solution to this problem is to distinguish between implicit and explicit knowledge and to construe the knowledge operator as capturing the concept of implicit knowledge. In addition, since a proposition is said to be implicitly known just in case it is derivable from the set of propositions that are explicitly known by using a certain set of logical rules, the concept of implicit knowledge is definable on the basis of the concept of explicit knowledge. In any case, both implicit and explicit knowledge are typically characterized as factive, i.e. such that it is always the case that what is known is also true. The aim of the present paper is twofold: first, we will develop a dynamic system of explicit intersubjective knowledge that allows us to introduce the operator of implicit knowledge by definition; secondly, we will show that it is not possible to hold together the following two theses: 1) the concept of implicit knowledge is definable along the lines indicated above; 2) the concept of implicit knowledge is factive.

Keywords: epistemic logic; possible worlds semantics; explicit knowledge; implicit knowledge; logical omniscience; factivity; inference.

1. The concept of implicit knowledge

The lack of knowledge of an agent can be modelled by introducing possible alternative situations. Indeed, if we do not know whether *p* is the case, then we can both imagine that the actual world is a world in which *p* is the case and imagine that the actual world is a world in which *p* is not the case. In addition, if we are in a position in which it is impossible to exclude either that the world is such that *p* is the case or that the world is such that *p* is not the case, then we lack knowledge about *p*. As a consequence, the knowledge of an agent can be modelled by introducing possible alternative situations and by assuming that an agent knows that *p* just in case all the situations that are not excluded by the agent are situations where *p* is the case. This intuition provides the basis to introduce the current modal definition of the knowledge operator[1].

1.1. Systems of implicit knowledge

The previous idea can be made precise by using a possible worlds semantics. Let *P* be a set of propositional variables. The set $L(P,\boldsymbol{K})$ of epistemic formulas is inductively defined according to the following rules:

$$\varphi := p \mid \neg\varphi \mid \varphi\wedge\varphi' \mid \boldsymbol{K}(\varphi), \text{ where } p\in P.$$

[1] See [8] and [10] for standard introductions to this topic.

The other propositional connectives are defined in the usual way. Intuitively, a formula like $\boldsymbol{K}(\varphi)$ says that the epistemic agent knows that φ.

Definition 1.1: frame.
A frame for $L(P,\boldsymbol{K})$ is a pair $F = (W,R)$, where

i) W is a non-empty set of worlds
ii) $R \subseteq W \times W$ is the accessibility relation on W.

Definition 1.2: model.
A model for $L(P,\boldsymbol{K})$ is a pair $M = (F,V)$, where

i) F is a frame for $L(P,\boldsymbol{K})$
ii) $V: P \to \wp(W)$ is a modal valuation on W.

Definition 1.3: truth at a world in a model ($M,w \models \varphi$).
Truth at a world in a model is inductively defined as follows:

$M,w \models p \Leftrightarrow w \in V(p)$
$M,w \models \neg\varphi \Leftrightarrow$ not $M,w \models \varphi$
$M,w \models \varphi\wedge\varphi' \Leftrightarrow M,w \models \varphi$ and $M,w \models \varphi'$
$M,w \models \boldsymbol{K}(\varphi) \Leftrightarrow \forall v \in W(R(w,v) \Rightarrow M,v \models \varphi)$

If φ is a formula, $\Vdash \varphi$ states that φ is valid in the class of all frames.

The definition of frame does not put any condition on R. By imposing conditions on R it is possible to get additional properties of knowledge. These properties are then reflected by systems of axioms that describe the valid formulas in classes of frames that satisfy the conditions. It is well-known that the following axioms characterize the conditions:

T: $\boldsymbol{K}(\varphi) \to \varphi$	R reflexive
4: $\boldsymbol{K}(\varphi) \to \boldsymbol{K}(\boldsymbol{K}(\varphi))$	R transitive
5: $\neg\boldsymbol{K}(\varphi) \to \boldsymbol{K}(\neg\boldsymbol{K}(\varphi))$	R euclidean

It is generally acknowledged that a suitable system of knowledge should lie between the system ***KT4*** and the system ***KT5***. Indeed, it is a distinctive characteristic of knowledge that the proposition that is known is also true, as stated by axiom **T**, and it is commonly admitted that knowledge is subject to a certain kind of introspection, so to allow the introduction of axiom **4**, stating that what is known is also known to be known.

1.2. The problem of logical omniscience

The foregoing characterization is subjected to the logical omniscience problem. In particular, it is not difficult to show that each instance of the following rules turns out to hold:

Omni1: $\Vdash \varphi \Rightarrow \Vdash \boldsymbol{K}(\varphi)$
Omni2: $\Vdash \varphi \rightarrow \varphi' \Rightarrow \Vdash \boldsymbol{K}(\varphi) \rightarrow \boldsymbol{K}(\varphi')$
Omni3: $\Vdash \varphi \leftrightarrow \varphi' \Rightarrow \Vdash \boldsymbol{K}(\varphi) \leftrightarrow \boldsymbol{K}(\varphi')$
Omni4: $\Vdash \boldsymbol{K}(\varphi \rightarrow \varphi') \rightarrow (\boldsymbol{K}(\varphi) \rightarrow \boldsymbol{K}(\varphi'))$

It is also evident that even a moderately idealized epistemic agent is incapable of knowing all the valid propositions, all the consequence of what she knows, and all the propositions that are logically equivalent to the propositions she knows. Hence, the concept represented by the operator ***K*** is not identified with the common concept of knowledge and the classical solution to the problem of identifying the concept of knowledge represented by ***K*** is to distinguish between implicit and explicit knowledge. In this way, the modal operator can be construed as capturing the concept of implicit knowledge, where implicit knowledge is introduced as the epistemic disposition to assent to all the consequences of what is explicitly known.[2] Still, once this distinction is introduced, it becomes possible to try to model an operator of explicit knowledge directly and to introduce the operator of implicit knowledge by definition, so that a better understanding of the logical characteristics of this kind of knowledge becomes available.
In what follows we will develop a logical system for explicit ad implicit intersubjective knowledge (section 2), show that such a system is sound and complete with respect to a suitable semantics (section 3), and show that within such a system factivity of implicit knowledge is not derivable (section 4). Finally, we will discuss a second way of coping with the problem of logical omniscience and conclude that factivity can be obtained only at the cost of a huge amount of idealization (section 5).

2. The concept of explicit intersubjective knowledge

To define an appropriate concept of explicit knowledge is a difficult task. One way to accomplish it is to model explicit knowledge by

1) limiting the set of inferential steps the agent can do;
2) limiting the set of inferential rules the agent can apply;
3) limiting the set of propositional contents the agent can handle.

[2] See [7], [8] and [12] for an introduction. In [7], pp 317-8, we find: "To represent the knowledge of agent *i*, we allow two modal operators K_i and X_i, standing for *implicit knowledge* and *explicit knowledge* of agent *i*, respectively. Implicit knowledge is the notion we have been considering up to now: truth in all worlds that the agent considers possible. On the other hand, an agent explicitly knows a formula φ if he is aware of φ and implicitly knows φ. Intuitively, an agent's implicit knowledge includes all the logical consequences of his explicit knowledge". A similar characterization is given in [4], p.11, and [14], pp. 241-2.

Let C_R be the operator of logical closure relative to a certain set *R* of rules. Let C_R^n be the operator that returns the set of propositions obtained by *n* successive applications of all the rules in *R* to a set of propositions. Once C_R^n is defined, we can constrain the inferential power of an agent by simply choosing a limited set of rules and a bound for the number of inferential steps the agent can take. Hence, we can define explicit knowledge by stating that a proposition is explicitly known just in case it is in $C_R^n(X)$ for a certain set *X* of initial propositions.[3] Still, even in this case, the agent turns out to be a perfect, though limited, knower, since she gets:

1) perfect knowledge relative to a limited set of inferential rules.
2) perfect knowledge relative to a limited set of inferential steps.
3) perfect knowledge relative to a limited set of propositional contents.

As a consequence, such an approach is not succeeding in representing the explicit knowledge of an actual epistemic agent.

A suitable strategy to overcome this difficulty is the one adopted and developed in [5] and [6]. The basic idea is that, if an agent explicitly knows all the propositions of a certain set and is explicitly aware of all the rules of a certain kind, then she is able to progressively deduce all the conclusions that are implicit in the initial propositions on the basis of the available rules. This idea can be implemented by making epistemic logic dynamic. Intuitively, an agent is viewed as a dynamical system capable of improving the set of known propositions, where the improvement is accomplished by steps, each step consisting in applying a deduction rule to some known propositions. If the aim is to characterize implicit knowledge as knowledge concerning what is deducible from explicit knowledge, then both the number of steps and the rules according to which the steps are accomplished are not to be bounded. Thus, the best way to construct such a dynamical system is to introduce a unique modal operator modelling the fact that a proposition is true after any inferential process accomplished by the agent. This operator corresponds to an accessibility relation linking a possible world *w* to a world *v* just in case the set of known propositions at *w* is deductively improved at *v*.[4]

A final consideration on the concept of explicit *intersubjective* knowledge is in order: intersubjective knowledge is characterized by the assumption that it concerns stable

[3] This kind of approach is proposed in [13]. See [1] and [11] for further developments.

[4] The system introduced in the following section is a modification of the system ***DES4*** of dynamic epistemic logic described and discussed in [6]. ***DES4*** was introduced without an appropriate semantics, but a Kripke style semantics was then presented in [1]. This system has some shortcomings, due to the fact that not every formula of the language in which the system is formulated is implicitly knowable: in particular, not every axiom is implicitly knowable. This limitation is counterintuitive and hinders a direct interpretation of implicit knowledge as knowledge deriving from a possible chain of inferential steps, since every axiom is surely accessible in a unique inferential step. However, it is worth noticing that the system we are going to introduce is more powerful than ***DES4***, so that the limitative conclusions we will prove are valid with respect to ***DES4*** as well. The present modification originates in the works about explicit logic of knowledge presented in [2] and [9]. Actually, the operator of implicit knowledge can be viewed as the existential generalization of the explicit operators adopted in these papers.

propositions, i.e. propositions that cannot change their truth value over time. The basic idea is that propositional knowledge is intersubjective when it is intersubjectively assessable, i.e. assessable by a scientific community, and this implies being dependably communicable, e.g. being communicable through a scientific journal. Thus, the content of intersubjective knowledge is not only true, but also cumulative in principle. In this sense, intersubjectively knowable propositions have to be expressed by sentences whose truth value is time invariant. To be sure, in order to be intersubjectively assessed, and so known, the truth value of a sentence has not to be relative to the situation in which the sentence is uttered, since that situation (i) could not be present to the agent that assesses the proposition and (ii) could change over time, so that our knowledge would not be communicable in a dependable way, since the time we spend in communicating it could be in principle sufficient for its truth value to change. Thus, all the information depending on the situation in which the sentence is uttered, e.g. the information about the time and the context of the utterance, has to be incorporated into the sentence itself, so to produce a sentence whose truth value is time invariant. In conclusion, a sentence expressing an intersubjectively knowable propositions is a sentence whose truth value is independent of the situations in which it is uttered, i.e. a sentence expressing a complete, and so stable, proposition.

In what follows, we will exploit this assumption in a limited version only, according to which intersubjective knowledge is about propositions whose truth values persist through any inferential step an agent can take.

2.1. The language of dynamic epistemic logic.
The language L of dynamic epistemic logic is defined as follows.

$$\varphi := p \mid \neg\varphi \mid \varphi\wedge\varphi' \mid \mathbf{k}(\varphi) \mid [\mathbf{F}]\varphi, \text{ where } p \text{ is in } P.$$

The other connectives are introduced according to the usual definitions.
$\mathbf{k}(\varphi)$ is interpreted as stating that φ is explicitly known, while $[\mathbf{F}]\varphi$ is interpreted as stating that φ is true after any inferential course, where an inferential course is conceived of as a sequence of basic inferential steps. In addition, we will use $\langle\mathbf{F}\rangle$ as the dual of $[\mathbf{F}]$, so that $\langle\mathbf{F}\rangle\varphi$ says that φ is true after at least one inferential course. As a consequence, a proposition like $\langle\mathbf{F}\rangle\mathbf{k}(\varphi)$ is interpreted as stating that φ is explicitly known after a certain number of inferential steps. We can then introduce the following

Definition 2.1: *explicit knowledge*.
A formula φ is explicitly known := $\mathbf{k}(\varphi)$.

Definition 2.2: *implicit knowledge*.
A formula φ is implicitly known := $\langle\mathbf{F}\rangle\mathbf{k}(\varphi)$.

2.2. The basic system

The basic system ***DE*** of dynamic epistemic logic for explicit and implicit intersubjective knowledge consists of three groups of axioms.

Group I: propositional axioms and rules:
all propositionally valid formulas
$\vdash \varphi$ and $\vdash \varphi \rightarrow \varphi'$ => $\vdash \varphi'$.

Group II: modal axioms and rules:
F1: $[\mathbf{F}](\varphi \rightarrow \varphi') \wedge [\mathbf{F}]\varphi \rightarrow [\mathbf{F}]\varphi'$
F2: $\langle \mathbf{F} \rangle \langle \mathbf{F} \rangle \varphi \rightarrow \langle \mathbf{F} \rangle \varphi$
F3: $\varphi \rightarrow \langle \mathbf{F} \rangle \varphi$
RF: $\vdash \varphi$ => $\vdash [\mathbf{F}]\varphi$

Group III: mixed axioms and rules:
FK1: $\langle \mathbf{F} \rangle \mathbf{k}(\varphi \rightarrow \varphi') \wedge \langle \mathbf{F} \rangle \mathbf{k}(\varphi) \rightarrow \langle \mathbf{F} \rangle \mathbf{k}(\varphi')$
FK2: $\mathbf{k}(\varphi) \rightarrow [\mathbf{F}]\varphi$
FK3: $\mathbf{k}(\varphi) \rightarrow \langle \mathbf{F} \rangle \mathbf{k}(\mathbf{k}(\varphi))$
RFK: $\vdash \langle \mathbf{F} \rangle \mathbf{k}(\varphi)$, provided φ is in *AX* = the set of axioms in Groups I-III.

Axioms of Group II characterize [**F**]. Hence, on the assumption that there are steps in which no inference is performed and that sequences of inferential steps can be chained, axioms **F2** and **F3** are justified.

Axioms of Group III characterize **k**. **FK1** states that, if the agent is able to perform both an inferential course that ends with the knowledge of $\varphi \rightarrow \varphi'$ and an inferential course that ends with the knowledge of φ, then she is always able to perform an inferential course that ends with the knowledge of φ'. Thus, an agent is assumed to be always able to store the conclusions obtained by performing different inferential courses and develop new inferences on the basis of them. **FK2** introduces the crucial condition to the effect that knowledge concerns *stable* propositions, i.e. that it is intersubjective. Finally **FK3** states that knowledge can be acknowledged, while **RFK** states that all axiom instances are knowable.[5]

Proposition 1: the following theorems are derivable.

[5] As a referee pointed out, a consequence of axiom **FK2** is that, if a proposition φ is knowable, $\langle \mathbf{F} \rangle \mathbf{k}(\varphi)$, then it is not possible to explicitly know that φ is not explicitly known, $\mathbf{k}\neg\mathbf{k}(\varphi)$. This consequence is acceptable insofar as the concept of knowledge captured by **k** is the concept of intersubjective and stable knowledge. Indeed, knowledge of $\neg\mathbf{k}(\varphi)$ is not stable, provided that φ is assumed to be knowable.

1.1: $\vdash_{\mathbf{DE}}$ $\mathbf{k}(\varphi) \rightarrow \varphi$.

Straightforward, by **F3** and **FK2**.

1.2: $\vdash_{\mathbf{DE}}$ $\mathbf{k}(\varphi) \rightarrow \langle\mathbf{F}\rangle\mathbf{k}\varphi$.

Straightforward, by **F3**.

1.3: $\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi) \rightarrow \langle\mathbf{F}\rangle\mathbf{k}(\mathbf{k}(\varphi))$.

$\vdash_{\mathbf{DE}}$ $\mathbf{k}(\varphi) \rightarrow \langle\mathbf{F}\rangle\mathbf{k}(\mathbf{k}(\varphi))$, by **FK3**.
$\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi) \rightarrow \langle\mathbf{F}\rangle\langle\mathbf{F}\rangle\mathbf{k}(\mathbf{k}(\varphi))$, by **RF** and **F1**.
$\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi) \rightarrow \langle\mathbf{F}\rangle\mathbf{k}(\mathbf{k}(\varphi))$, by **F2**.

1.4: $\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi) \rightarrow \langle\mathbf{F}\rangle\mathbf{k}(\langle\mathbf{F}\rangle\mathbf{k}(\varphi))$.

$\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\mathbf{k}(\varphi) \rightarrow \langle\mathbf{F}\rangle\mathbf{k}(\varphi))$, by **RFK** from **F3**.
$\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\mathbf{k}(\varphi)) \rightarrow \langle\mathbf{F}\rangle\mathbf{k}(\langle\mathbf{F}\rangle\mathbf{k}(\varphi))$, by **FK1**.
$\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi) \rightarrow \langle\mathbf{F}\rangle\mathbf{k}(\langle\mathbf{F}\rangle\mathbf{k}(\varphi))$, by 1.3.

1.5: $\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi) \rightarrow \langle\mathbf{F}\rangle\mathbf{k}([\mathbf{F}]\varphi)$.

$\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\mathbf{k}(\varphi) \rightarrow [\mathbf{F}]\varphi)$, by **RFK** from **FK2**.
$\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\mathbf{k}(\varphi)) \rightarrow \langle\mathbf{F}\rangle\mathbf{k}([\mathbf{F}]\varphi))$, by **FK1**.
$\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi) \rightarrow \langle\mathbf{F}\rangle\mathbf{k}([\mathbf{F}]\varphi)$, by 1.3.

Proposition 2: internalization.
RFK!: $\vdash_{\mathbf{DE}}$ φ => $\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi)$.
The proof is by induction on the length of derivations in **DS4**.

Base: φ is an axiom instance.
The conclusion follows by **RFK**.

Step 1: φ is derived from φ' and $\varphi'\rightarrow\varphi$ by *modus ponens*.
By induction hypothesis, $\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi')$ and $\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi'\rightarrow\varphi)$.
Hence $\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi)$, by **FK1**.

Step 2: φ has the form $[\mathbf{F}]\varphi'$ and is derived from φ' by **RF**.
By induction hypothesis, $\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}(\varphi')$.
Hence $\vdash_{\mathbf{DE}}$ $\langle\mathbf{F}\rangle\mathbf{k}([\mathbf{F}]\varphi')$, by proposition 1.5.

Step 3: φ has the form $\langle \mathbf{F} \rangle \mathbf{k}(\varphi')$ and is derived from φ' by **RFK**.
By induction hypothesis, $\vdash_{\mathbf{DE}} \langle \mathbf{F} \rangle \mathbf{k}(\varphi')$.
Hence $\vdash_{\mathbf{DE}} \langle \mathbf{F} \rangle \mathbf{k}(\langle \mathbf{F} \rangle \mathbf{k}(\varphi'))$, by proposition 1.4.

Proposition 3: the following rules are derivable.
RFKI: $\vdash_{\mathbf{DE}} \varphi \rightarrow \varphi' \Rightarrow \vdash_{\mathbf{DE}} \langle \mathbf{F} \rangle \mathbf{k}(\varphi) \rightarrow \langle \mathbf{F} \rangle \mathbf{k}(\varphi')$.
RFKE: $\vdash_{\mathbf{DE}} \varphi \leftrightarrow \varphi' \Rightarrow \vdash_{\mathbf{DE}} \langle \mathbf{F} \rangle \mathbf{k}(\varphi) \leftrightarrow \langle \mathbf{F} \rangle \mathbf{k}(\varphi')$.
Straightforward, by proposition 2 and **FK1**.

RFK! is a crucial rule, since it declares that the set of derivable formulas is closed with respect to the operator of implicit knowledge. Hence, **RFK!** corresponds to the rule **Omni1** above. In addition, **RFKI**, **RFKE** and **FK1** reflect **Omni2**, **Omni3** and **Omni4**, so that, in accordance with our expectations, all the rules on omniscience are valid with respect to the concept of implicit knowledge.[6]

Proposition 4 (*consistency*): $\vdash_{\mathbf{DE}} \langle \mathbf{F} \rangle \mathbf{k}(\varphi) \rightarrow \neg \langle \mathbf{F} \rangle \mathbf{k} \neg(\varphi)$.
$\vdash_{\mathbf{DE}} \langle \mathbf{F} \rangle \mathbf{k}(\varphi) \wedge \langle \mathbf{F} \rangle \mathbf{k}(\neg\varphi) \rightarrow \langle \mathbf{F} \rangle \mathbf{k}(\varphi \wedge \neg\varphi)$, by **RFK!** and **FK1**.
$\vdash_{\mathbf{DE}} \mathbf{k}(\varphi \wedge \neg\varphi) \rightarrow (\varphi \wedge \neg\varphi)$, by **FK2** and **F3**.
$\vdash_{\mathbf{DE}} \langle \mathbf{F} \rangle \mathbf{k}(\varphi \wedge \neg\varphi) \rightarrow \langle \mathbf{F} \rangle (\varphi \wedge \neg\varphi)$, by **RFK!** and **FK1**.
$\vdash_{\mathbf{DE}} \neg \langle \mathbf{F} \rangle (\varphi \wedge \neg\varphi)$, by **RF**.
$\vdash_{\mathbf{DE}} \neg \langle \mathbf{F} \rangle \mathbf{k}(\varphi \wedge \neg\varphi)$, by propositional logic.
$\vdash_{\mathbf{DE}} \langle \mathbf{F} \rangle \mathbf{k}(\varphi) \rightarrow \neg \langle \mathbf{F} \rangle \mathbf{k} \neg(\varphi)$, by propositional logic.

Proposition 4 states the intuitive principle that implicit knowledge is consistent.

2.3. Note on the basic system

The system introduced above in based on a syntactic approach, since explicit knowledge is modelled on the basis of sets of formulas assigned to possible worlds, and allows us[7]

(1) to construe implicit knowledge in terms of explicit knowledge
(2) to interpret a standard ***KT4*** system of implicit knowledge
(3) to use nested epistemic operators

It is worth noting that the internalization theorem is necessary to interpret a standard ***KT4*** system of implicit knowledge and that the mixed axioms and rules in **Group III**

[6] As suggested by a referee, we should say that **RFK!** and the other rules and axioms corresponds to the dynamic versions of the omniscience principles. Still, if one accepts the interpretation of ***K*** in terms of $\langle \mathbf{F} \rangle \mathbf{k}$, the correspondence is complete.

[7] As suggested by a referee, the approach we follow is probably the only one which provides us with (1)-(3). Indeed, the current syntactic approaches either introduce the explicit operators only as primitive, but do not treat nested operators (see [1], [5], [6], [11]), or treat nested operators, but introduce implicit and explicit operators as primitive (see [7], [15], [16], [17]).

are necessary in order to prove the internalization theorem. This is why we have chosen the present system as the basic one. Still, the limitations we are going to display affect any system that is intermediate between our basic system and the very weak one determined by axioms in Group I, Group II and the following Group III*.

Group III*: mixed axioms and rules:
FK1*: $\mathbf{k}(\varphi \rightarrow \varphi') \wedge \mathbf{k}(\varphi) \rightarrow \langle \mathbf{F} \rangle \mathbf{k}(\varphi')$
FK2*: $\mathbf{k}(\varphi) \rightarrow \varphi$, where φ has no occurrences of [**F**] or **k**.
FK3*: $\mathbf{k}(\varphi) \rightarrow \langle \mathbf{F} \rangle \mathbf{k}(\mathbf{k}(\varphi))$, where φ has no occurrences of [**F**] or **k**.
RFK*: $\vdash \langle \mathbf{F} \rangle \mathbf{k}(\varphi)$, provided φ is in AX = the set of axioms in group I + **FK2***.

3. Semantics

The basic idea underlying the semantics for the present system of explicit and implicit knowledge is to model explicit knowledge by associating with each world *w* the set of formulas that are explicitly known at *w* and then to use definition 2.2 in order to introduce implicit knowledge.[8]

Definition 3.1: frame.
A frame for $L(\boldsymbol{DE})$ is a triple $F = (W,R,K)$ where

1) W is a non-empty set.
2) $R \subseteq W \times W$ is a pre-order on W.
3) K: $W \rightarrow \wp(L(\boldsymbol{DE}))$ is a selection function on W.

Intuitively, a selection function K assigns to each world the set of formulas that are explicitly known by the agent at that world. In addition, R and K have to respect the following conditions (where $K^*(w) = \cup\{K(v) \mid R(w,v)\}$).

K1): $\varphi, \varphi \rightarrow \varphi' \in K^*(w) \Rightarrow \varphi' \in K^*(w)$
K2): $\varphi \in K(w) \Rightarrow \mathbf{k}(\varphi) \in K^*(w)$
KAX): $\varphi \in AX \Rightarrow \varphi \in K^*(w)$

Definition 3.2: model.
A model for $L(\boldsymbol{DE})$ is a pair $M = (F,V)$, where

i) F is a frame for $L(\boldsymbol{DE})$
ii) V: $P \rightarrow \wp(W)$ is a modal valuation on W.

[8] The idea of assigning set of formulas to worlds in order to model what is explicitly known is proposed in [7] and developed in [8], [9] and [2] in different directions.

Definition 3.3: truth at a world in a model ($M,w \models \varphi$).
Truth at a world in a model is inductively defined according to the following rules:

$M,w \models p \iff w \in V(p)$
$M,w \models \neg\varphi \iff$ not $M,w \models \varphi$
$M,w \models \varphi \wedge \varphi' \iff M,w \models \varphi$ and $M,w \models \varphi'$
$M,w \models [\mathbf{F}]\varphi \iff \forall v \in W(R(w,v) \Rightarrow M,v \models \varphi)$
$M,w \models \mathbf{k}(\varphi) \iff \varphi \in K(w)$

Notice that $M,w \models \langle \mathbf{F} \rangle \mathbf{k}(\varphi) \iff \exists v(R(w,v)$ and $\varphi \in K(v)) \iff \varphi \in K^*(w)$.

Definition 3.4: admissible model.
An admissible model for $L(\boldsymbol{DE})$ is a model M satisfying

KAD): $\varphi \in K(w)$ and $R(w,v) \Rightarrow M,v \models \varphi$

Hence, an admissible model is a model in which the propositions that are known at a given world w are true at every world accessible to w, in particular at w itself.

Theorem 1: ***DE*** is sound with respect to the class of all admissible models.
Proof. Soundness is proved by induction on the length of derivations. The proof of the validity of the axioms and rules of the first two groups is standard, while the validity of the axioms and rules of the third group is a straight consequence of conditions *K1*-*KAX* and *KAD*. We only check the validity of **FK1**.

FK1: $\Vdash \langle \mathbf{F} \rangle \mathbf{k}(\varphi \rightarrow \varphi') \wedge \langle \mathbf{F} \rangle \mathbf{k}(\varphi) \rightarrow \langle \mathbf{F} \rangle \mathbf{k}(\varphi')$.
Suppose $M,w \models \langle \mathbf{F} \rangle \mathbf{k}(\varphi \rightarrow \varphi')$ and $M,w \models \langle \mathbf{F} \rangle \mathbf{k}(\varphi)$.
Then $\varphi \rightarrow \varphi' \in K^*(w)$ and $\varphi \in K^*(w)$. Thus $\varphi' \in K^*(w)$, by *K1*, and so $M,w \models \langle \mathbf{F} \rangle \mathbf{k}(\varphi')$.

Theorem 2: ***DE*** is complete with respect to the class of all admissible models.
Proof. Completeness is proved by canonicity. If X is a ***DE*** consistent set of formulas, then X can be extended to a maximal consistent set x by a standard procedure[9]. In what follows w, v, and so on, will range over maximal ***DE*** consistent sets.

Definition 3.5: canonical model.
Let $w/[\mathbf{F}] = \{\varphi \mid [\mathbf{F}]\varphi \in w\}$ and $w/\mathbf{k} = \{\varphi \mid \mathbf{k}(\varphi) \in w\}$.
The canonical model is the tuple $\langle W,R,K,V \rangle$, where

– W is the set of maximal ***DE*** consistent sets.

[9] The standard modal parts of the proof of completeness are omitted. See [3], ch. 4, for details.

– R is such that $R(w,v)$ <=> $w/[\mathbf{F}] \subseteq v$.
– K is such that $K(w) = w/\mathbf{k}$.
– V is such that $V(w) = \{p \mid p\in w\}$.

Fact 1: $w/[\mathbf{F}]$ is a ***DE*** closed set.
Suppose $w/[\mathbf{F}] \vdash_{DE} \varphi$. Then $w \vdash_{DE} [\mathbf{F}]\varphi$, by the definition of $w/[\mathbf{F}]$. Thus $[\mathbf{F}]\varphi\in w$, since w is maximal consistent, and so $\varphi\in w/[\mathbf{F}]$, again by the definition of $w/[\mathbf{F}]$.

Fact 2: $\langle\mathbf{F}\rangle\mathbf{k}(\varphi)\in w$ <=> $\exists v(w/[\mathbf{F}] \subseteq v$ and $\varphi\in v/\mathbf{k})$.
$\langle\mathbf{F}\rangle\mathbf{k}(\varphi)\in w$ <=> $[\mathbf{F}]\neg\mathbf{k}(\varphi)\notin w$, since w is complete
$\langle\mathbf{F}\rangle\mathbf{k}(\varphi)\in w$ <=> $\neg\mathbf{k}(\varphi)\notin w/[\mathbf{F}]$, by the definition of $w/[\mathbf{F}]$
$\langle\mathbf{F}\rangle\mathbf{k}(\varphi)\in w$ <=> not $w/[\mathbf{F}] \vdash_{DE} \neg\mathbf{k}(\varphi)$, since $w/[\mathbf{F}]$ is closed
$\langle\mathbf{F}\rangle\mathbf{k}(\varphi)\in w$ <=> $\exists v(w/[\mathbf{F}] \subseteq v$ and $\mathbf{k}(\varphi)\in v)$, since $w/[\mathbf{F}]\cup\{\mathbf{k}(\varphi)\}$ is consistent
$\langle\mathbf{F}\rangle\mathbf{k}(\varphi)\in w$ <=> $\exists v(w/[\mathbf{F}] \subseteq v$ and $\varphi\in v/\mathbf{k})$, by the definition of $w/[\mathbf{F}]$

Lemma 1 (*truth lemma*): for all w, $M,w \models \varphi$ <=> $\varphi\in w$.
The only non standard case is when a formula has the form $\mathbf{k}(\varphi)$.
In this case: $M,w \models \varphi$ <=> $\varphi\in K(w)$ <=> $\varphi\in w/\mathbf{k}$ <=> $\mathbf{k}(\varphi)\in w$.

Lemma 2: $\langle W,R,K,V\rangle$ is a model for ***DE***.
R is a pre-order.
Standard, by **F2** and **F3**.

K1): $\varphi,\varphi\rightarrow\varphi'\in K^*(w)$ => $\varphi'\in K^*(w)$.
By fact 2, it suffices to prove that
$\langle\mathbf{F}\rangle\mathbf{k}(\varphi\rightarrow\varphi')\in w$ and $\langle\mathbf{F}\rangle\mathbf{k}(\varphi)\in w$ => $\langle\mathbf{F}\rangle\mathbf{k}(\varphi')\in w$.
Suppose $\langle\mathbf{F}\rangle\mathbf{k}(\varphi\rightarrow\varphi')\in w$ and $\langle\mathbf{F}\rangle\mathbf{k}(\varphi)\in w$. Then $\langle\mathbf{F}\rangle\mathbf{k}(\varphi')\in w$, by **FK1**.

K2): $\varphi\in K(w)$ => $\mathbf{k}(\varphi)\in K^*(w)$
By fact 2, it suffices to prove that
$\varphi\in w/\mathbf{k}$ => $\langle\mathbf{F}\rangle\mathbf{k}(\mathbf{k}(\varphi))\in w$.
Suppose $\mathbf{k}(\varphi)\in w$. The conclusion follows by **FK3**.

KAX): $\varphi\in AX$ => $\varphi\in K^*(w)$.
By fact 2, it suffices to prove that $\varphi\in AX$ => $\langle\mathbf{F}\rangle\mathbf{k}(\varphi)\in w$.
Suppose $\varphi\in AX$. Then $\vdash_{DE} \varphi$, and so $\langle\mathbf{F}\rangle\mathbf{k}(\varphi)\in w$, by **RFK**.

Lemma 3: $\langle W,R,K,V\rangle$ is an admissible model for ***DE***.
KAD): $\varphi\in K(w)$ and $R(w,v)$ => $M,v \models \varphi$
By lemma 1 and fact 2, it suffices to prove that
$\varphi\in w/\mathbf{k}$ and $w/[\mathbf{F}] \subseteq v$ => $\varphi\in w$, i.e., $\mathbf{k}(\varphi)\in w$ and $w/[\mathbf{F}] \subseteq v$ => $\varphi\in v$.

Suppose $\mathbf{k}(\varphi)\in w$. Then $[\mathbf{F}]\varphi\in w$, by **FK2**, and so $\varphi\in v$, since $w/[\mathbf{F}] \subseteq v$.

This concludes the proof.

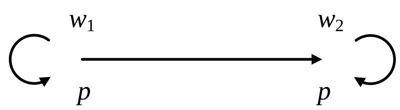
4. Limitations

In this section we show that in ***DE*** the concepts of implicit and explicit knowledge are different. In particular, in ***DE*** it is not possible to derive $\langle\mathbf{F}\rangle\mathbf{k}(\varphi) \rightarrow \mathbf{k}(\varphi)$. Moreover, in ***DE*** it is not possible to derive $\langle\mathbf{F}\rangle\mathbf{k}(\varphi) \rightarrow \varphi$.

Theorem 3: $\langle\mathbf{F}\rangle\mathbf{k}(\varphi) \rightarrow \mathbf{k}(\varphi)$ is not derivable in ***DE***.
In order to prove it, let us introduce the following model.

– $W = \{w_1,w_2\}$
– $R = \{\langle w_1,w_1\rangle,\langle w_1,w_2\rangle,\langle w_2,w_2\rangle\}$
– K is such that $K(w_1) = \varnothing$ and $K(w_2)$ = the smallest set X such that
    – $AX \subseteq X$
    – $\varphi,\varphi\rightarrow\varphi'\in X \Rightarrow \varphi'\in X$
    – $\varphi\in X \Rightarrow [\mathbf{F}]\varphi\in X$ and $\mathbf{k}(\varphi)\in X$
– V is such that $V(p) = W$ and $V(p') = \varnothing$, for p' different from p.

In a picture:



We first show that $\langle W,R,K,V\rangle$ is an adequate model for ***DE***. It is plain that R is a pre-order on W and that K: $W \rightarrow \wp(L(\boldsymbol{DE}))$. In addition, R and K respect *K1*)-*KAX*) and the condition *KAD*). First of all, notice that $K(w_1) = \varnothing$ and $K^*(w_1) = K^*(w_2) = K(w_2)$, by the definition of K^*.

K1). $\varphi,\varphi\rightarrow\varphi'\in K^*(w) \Rightarrow \varphi'\in K^*(w)$.
Suppose $\varphi,\varphi\rightarrow\varphi'\in K^*(w)$, where $w\in\{w_1,w_2\}$.
Then $\varphi,\varphi\rightarrow\varphi'\in K(w_2)$, and so $\varphi'\in K(w_2)$, by the definition of $K(w_2)$.
Thus, $\varphi'\in K^*(w)$, where $w\in\{w_1,w_2\}$, since $K^*(w_1) = K^*(w_2) = K(w_2)$.

K2). $\varphi\in K(w) \Rightarrow \mathbf{k}(\varphi)\in K^*(w)$.
Suppose $\varphi\in K(w)$, where $w\in\{w_1,w_2\}$.
If $w = w_1$, then the conclusion follows, since $K(w_1) = \varnothing$.
If $w = w_2$, the conclusion follows, since $\mathbf{k}(\varphi)\in K(w_2) = K^*(w_2)$, by the definition of K.

KAX): $\varphi\in AX \Rightarrow \varphi\in K^*(w)$.
Straightforward, since $AX \subseteq K(w_2) = K^*(w_1) = K^*(w_2)$.

KAD): $\varphi \in K(w)$ and $R(w,v)$ => $M,v \models \varphi$.
Suppose $\varphi \in K(w)$ and $R(w,v)$, where $w,v \in \{w_1,w_2\}$
If $w = w_1$, then the conclusion follows, since $K(w_1) = \emptyset$. If $w = w_2$, then the conclusion follows by induction on the definition of $K(w_2)$. Indeed, it is not difficult to see that every axiom instance in *AX* is verified at w_2 and that, if $M,w_2 \models \varphi$ and $M,w_2 \models \varphi \rightarrow \varphi'$, then $M,w_2 \models \varphi'$. Suppose, then, that $\varphi \in K(w_2)$, where φ coincides with $[\mathbf{F}]\varphi' \in K(w_2)$ for a given $\varphi' \in K(w_2)$. Then, by inductive hypothesis, $M,w_2 \models \varphi'$ and, since w_2 only accedes to itself, $M,w_2 \models [\mathbf{F}]\varphi'$. Suppose, finally, that $\varphi \in K(w_2)$, where φ coincides with $\mathbf{k}(\varphi') \in K(w_2)$ for a given $\varphi' \in K(w_2)$. Then, by truth definition, $M,w_2 \models \mathbf{k}(\varphi')$.

Conclusion: $M,w_1 \models \langle \mathbf{F} \rangle \mathbf{k}(p \rightarrow p)$ and $M,w_1 \models \neg \mathbf{k}(p \rightarrow p)$.
i) not $M,w_1 \models \mathbf{k}(p \rightarrow p)$, since $K(w_1) = \emptyset$.
ii) $M,w_2 \models \langle \mathbf{F} \rangle \mathbf{k}(p \rightarrow p)$, since $R(w,v)$ and $p \rightarrow p \in AX \subseteq K(w_2)$.

Theorem 4: $\langle \mathbf{F} \rangle \mathbf{k}(\varphi) \rightarrow \varphi$ is not derivable in ***DE***.
Suppose $\vdash_{DE} \langle \mathbf{F} \rangle \mathbf{k}(\varphi) \rightarrow \varphi$.
$\vdash_{DE} \langle \mathbf{F} \rangle \mathbf{k}(\mathbf{k}(p \rightarrow p)) \rightarrow \mathbf{k}(p \rightarrow p)$.
$\vdash_{DE} \langle \mathbf{F} \rangle \mathbf{k}(p \rightarrow p) \rightarrow \mathbf{k}(p \rightarrow p)$, by proposition 1.3
Hence, we obtain a contradiction by the previous theorem.

In conclusion, *implicit knowledge*, as previously defined, *is not generally factive*. This conclusion is noteworthy, since factivity is assumed in almost every system of implicit knowledge and the definition of implicit knowledge as knowledge concerning the consequences of what is explicitly known is a customary one.

5. Eliminating the distinction between implicit and explicit knowledge
Let us consider the extension ***DEI*** of ***DE*** obtained by adding the following axiom:

FKI: $\langle \mathbf{F} \rangle \mathbf{k}(\varphi) \rightarrow \mathbf{k}(\varphi)$

FKI (for *idealized* knower) captures the idea that the epistemic agent has infinite inferential power, in the sense that she knows all the propositions she is able to infer from what she currently knows. Thus, if a proposition is implicitly known, $\langle \mathbf{F} \rangle \mathbf{k}(\varphi)$, then it is also explicitly known, $\mathbf{k}(\varphi)$. The corresponding semantic condition is:

KI): $K^*(w) = K(w)$.

It is not difficult to show that the logic so obtained is both sound and complete with respect to the class of frames characterized by *KI*.[10]

The important point is that the implicit knowledge operator $\langle\mathbf{F}\rangle\mathbf{k}$ is now eliminable in favour of $\mathbf{k}$. Accordingly, implicit knowledge becomes factive, but only at the cost of idealizing human knowledge in such a way that the distinction between implicitly and explicitly known propositions is eliminated. Moreover, the standard ***KT4*** epistemic logic of implicit knowledge can be embedded in the fragment of ***DEI*** concerning the knowledge operator. To be sure, let us consider the following translation function *:

i) $p^* = p$
ii) $(\neg\varphi)^* = \neg\varphi^*$
iii) $(\varphi\wedge\varphi')^* = \varphi^* \wedge \varphi'^*$
iv) $(\boldsymbol{K}(\varphi))^* = \mathbf{k}(\varphi^*)$

Fact 4: $X \vdash_{\boldsymbol{KT4}} \varphi \Leftrightarrow X^* \vdash_{\boldsymbol{DIE}} \varphi^*$.

From left to right: straightforward, by **RFK!**, **FK1-FK3**, **FKI** and **F3**.

From right to left: suppose not $X \vdash_{\boldsymbol{KT4}} \varphi$. Then, since ***KT4*** is sound and complete with respect to the class of all the frames in which R is reflexive and transitive, there is a model $M = \langle W,R,V\rangle$, where R is reflexive and transitive, and a world w such that $M,w \models X$ and not $M,w \models \varphi$. In order to obtain our conclusion it is then sufficient to show that, given M, it is possible to produce a model M' and a world w' such that $M',w' \models X^*$ and not $M',w' \models \varphi^*$. Let M' be the model $\langle W',R',K',V'\rangle$ such that

1) $W' = W$.
2) $R' = R$, so that it is a pre-order.
3) K' is defined by $K'(w) = \{\varphi^* \mid M,w \models \boldsymbol{K}(\varphi)\}$.
4) $V' = V$.

Notice that $K'^*(w) = \cup\{K'(v) \mid R(w,v)\} = \cup\{\{\varphi \mid M,v \models \boldsymbol{K}(\varphi)\} \mid R(w,v)\}$. However, since $M,w \models \boldsymbol{K}(\varphi)$ and $R(w,v)$ implies $M,v \models \boldsymbol{K}(\varphi)$, we have that $K(w) \subseteq K^*(w)$. It is not difficult to prove that conditions *K1*), *K2*), *KAX*) are satisfied and a straightforward induction on the length of a formula shows that $M,w \models \varphi \Leftrightarrow M',w \models \varphi^*$. Indeed, the only interesting case is the modal one, which is treated as follows:

$M,w \models \boldsymbol{K}(\varphi) \Leftrightarrow \varphi^*\in K'(w) \Leftrightarrow M',w \models \mathbf{k}(\varphi^*) \Leftrightarrow M',w \models (\boldsymbol{K}(\varphi))^*$

[10] **FKI** provides us with a second strategy to copy with the problem of logical omniscience. As highlighted in [14], p. 242, there are two general strategies for facing the problem. According to a first one, we can extend the ordinary concept of belief by introducing the distinction between explicit and implicit belief. This is the strategy commonly proposed. According to a second one, we can extend the ordinary concept of agent, by idealizing her epistemic power along the lines proposed here.

Since $M,w \models \varphi \Leftrightarrow M',w \models \varphi^*$, for any formula φ, we have both that $M,w \models X$ if and only if $M',w \models X^*$ and that $M,w \models \varphi$ if and only if $M',w \models \varphi^*$, whence the conclusion.

A consequence of fact 4 is that the ***KT4*** system of implicit knowledge can be equated with the pure epistemic portion of the system of dynamic epistemic logic in which the distinction between implicit and explicit logic is neglected on the basis of the assumption that ascribes to the epistemic agent an infinite inferential power.

Conclusion

We have shown that it is possible to model the distinction between explicit and implicit knowledge in a consistent way and that in the proposed model implicit and intersubjective knowledge turns out to be non-factive. In particular, we have shown that, if implicit knowledge is conceived of in the usual way, i.e. in such a way that a proposition is implicitly known just in case it is a consequence of a set of explicitly known propositions, and implicit intersubjective knowledge is about stable propositions, then non-factivity is a consequence of the distinction between explicit and implicit knowledge and the assumption that knowledge is potentially introspective, i.e. such that, if we know φ, then we can infer the truth of **k**(φ). In conclusion, if we do not eliminate the distinction between implicit and explicit knowledge and the fact that the knowledge of φ implies the possibility to infer the truth of **k**(φ), then we are committed to the denial of the general validity of $\boldsymbol{K}(\varphi) \rightarrow \varphi$. Hence, since knowledge, in strict sense, is factive, this conclusion is to the effect that we are in need of an additional analysis toward an accurate logical characterization of the concept of knowledge. Such an analysis should conciliate, or constrain, (i) factivity, (ii) positive introspection and (iii) the definition of implicit intersubjective knowledge in terms of explicit intersubjective knowledge. As far as I can see, a preliminary step in this direction should consist in developing more explicit versions of the system ***DE***, based on the introduction of explicit proofs operators, thus following the intuitions in [2], [15], [16], [17] so to check the possibility of recovering factivity by limiting the operators that give rise to intersubjective knowledge.

References

[1] ÅGOTNES, T. and N. ALECHINA, The Dynamics of Syntactic Knowledge, Journal of Logic and Computation 17, 2007: 83–116.

[2] ARTEMOV, S., The Logic of Justification. The Review of Symbolic Logic, 1, 2008, 477-513.

[3] BLACKBURN, P., M. DE RIJKE, and Y. VENEMA, Modal Logic. Cambridge: Cambridge University Press 2001.

[4] CRESSWELL, J., Intensional Logics and Logical Truth, Journal of Philosophical Logic, 1, 1972: 2-15.
[5] DUC, H. N. Logical Omniscience vs. Logical Ignorance on a Dilemma of Epistemic Logic. Vol. 990 of Lecture Notes in Computer Science. pp. 237–248, Berlin – New York: Springer 1995.
[6] DUC, H. N., Reasoning about rational, but not logically omniscient, agents, Journal of Logic and Computation 7, 1997: 633–648.
[7] FAGIN, R. and HALPERN, J. Y., Belief, awareness, and limited reasoning, Artificial Intelligence, 34, 1988: 39-76.
[8] FAGIN, R., HALPERN, J., MOSES, Y., VARDI, M., Reasoning About Knowledge, Cambridge, MA: The MIT Press 1995.
[9] FITTING. M., The logic of proofs, semantically, Annals of Pure and Applied Logic, 132, 2005 :1–25.
[10] HOEK, W.v.d. and MEYER, J.J., Epistemic Logic for AI and Computer Science, Cambridge: CUP 1995.
[11] JAGO, M., Epistemic logic for rule-based agents. Journal of Logic, Language and Information, 18, 2009: 131–158.
[12] LEVESQUE, H.J., A logic of implicit and explicit belief, Proceedings AAAI-84, Austin, Texas, 1984: 198-202.
[13] KONOLIGE, K., A Deduction Model of Belief. S. Francisco: Morgan Kaufmann 1986.
[14] STALNAKER, R., Context and Content. Oxford: Oxford University Press 1999.
[15] VAN BENTHEM, J., Merging observation and access in dynamic logic. Journal of Logic Studies, 1, 2008: 1–17.
[16] VAN BENTHEM, J. and VELÁZQUEZ-QUESADA, F.R., The dynamics of awareness. Synthese 177 (Supplement-1): 2010: 5-27.
[17] VELÁZQUEZ-QUESADA, F.R. Inference and Update. Synthese 169, 2009: 283-300.