Unconscious Knowledge & Epistemic Accessibility

By LUÍS M. AUGUSTO

University of Sussex
DEPT OF PSYCHOLOGY
(Visiting Research Fellow)

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR
(Postdoctoral Fellow)

UNIVERSITY OF SUSSEX, DEC. 04, 2008

Research Funded by FCT & European Social Fund

Unconscious Knowledge

- The expression “unconscious knowledge” originates in experimental cognitive psychology, but can easily be applied in other cognate fields such as knowledge management, and, with some adjustments, psychoanalysis.
- We speak of unconscious knowledge when a subject *S* lacks ***metaknowledge*** concerning his/her epistemic states, i.e. s/he does not know that s/he knows (e.g.: Dienes & Perner, 2002). In other words, we say that *S* has unconscious knowledge when s/he has no ***epistemic access*** to part of his/her own knowledge base.
- Concrete examples of unconscious knowledge abound in experimental psychology; blindsight, left visuo-spatial neglect, prosopagnosia, experiments with artificial grammars, etc, provide us with such examples (Augusto, 2008).
- We can represent this lack of epistemic access by means of ‘ignorance sets’ such as **{*Kp, ¬KKp*}**, and **{*Bp, ¬BBp*}**.

Knowledge & Modal Logic

- We are interested in finding out the *conditions* in which a subject may be said to actually possess unconscious knowledge. These are the ***epistemic conditions*** that have to do with **knowledge** and **belief**. Modal logic is useful in that we can construct models of *learning*, *information*, and *strategies* that, on their turn, might help us to better understand these epistemic notions.
- Specifically, we want to know in which way modal logic might be of help concerning an understanding, at an epistemological level, of <lack of> epistemic access; we are justified in this expectation in that **epistemic accessibility** is a major component of epistemic logic, a subfield of modal logic.

Human Knowledge & Epistemic Logic

But logic deals with ***logical* inferences**, which are not necessarily true of human reasoning: for instance, that *S* knows *p* & *q* does not necessarily mean that *S* knows *p* and *q* separately. Thus, epistemic logic, the modal logic dealing with knowledge and belief, appears not to be of interest for human reasoning. However, *epistemology* is *not* concerned solely with human knowledge (What do we know?, How do we know?, etc), but with *knowledge per se*: **WHAT IS KNOWLEDGE?**

Therefore, epistemic logic seems to have a word to say on epistemological matters (de Bruin, 2007; Hendricks & Symons, 2006; van Benthem, 2006).

Namely, epistemic logic specifies the types of **access** one has to knowledge and, more interestingly, to one's own knowledge. This it does by means of the postulation of an ***accessibility relation R* between *possible worlds*** in an application of **Possible World Semantics** as an appropriate interpretation for epistemic logic systems.

Syntax: K

The term “modal logic” primarily means the logic that involves the use of the [modal] expressions ‘necessarily’ and ‘possibly’; however, it can also cover a family of logics with similar rules (deontic logic, temporal logic, and epistemic logic are the best known).

The vocabulary of modal logic is the same as that of classical logic augmented with the unary operations □ and ◊.

System **K**, the weakest normal modal logic, consists of the following **axioms**:

- Any axiomatization of propositional logic
- ***K*** axiom**:** $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$

It consists still of the following **rules**:

- (MP) *modus ponens*
- $p / \Box p$ (necessitation rule)

This system is *sound* (every theorem is a tautology) and *complete* (every tautology is a theorem) with respect to the class of all Kripke models = the set of theorems in this system is exactly the set of validities.

In order to obtain **stronger** logics, we must add further axioms:

T, S4, and S5

We obtain system **T** by *extending* system **K** with <u>axiom T</u>

$$\Box p \rightarrow p$$

By extending system **T** with the <u>axiom 4</u>

$$\Box p \rightarrow \Box\Box p$$

we obtain system **S4.**

<u>Axiom 5</u>

$$\Diamond p \rightarrow \Box\Diamond p$$

added to **S4** gives us system **S5.**

Semantics

- A Kripke structure *M* for *n* agents over $\Phi = \{P, Q, R, \ldots\}$ is a tuple *(W*, *R*, *v)*, where *W* is a nonempty set of *states* or *possible worlds*, *R* is a binary *accessibility relation* between possible worlds, and *v* is an *interpretation* which associates with each state in *W* a truth assignment to the primitive propositions in Φ.
- The truth assignment tells us whether or not a formula *p* is true or false in a certain state. So *v*(*w*)(*p*) tells us whether *p* is true in world *w* in model *M*. Truth depends not only on the structure, but on the current world as well. Just because something is true in a world does not mean it is true in another. To show that a formula is true in a certain world, one writes , normally $(M, w) \vDash p$; this reads as "*p* is true at (*M,w*)," or "(*M, w)* satisfies *p*."

Possible World Semantics & Accessibility

PWS introduces ***possible worlds*** and a binary ***accessibility relation R*** between pairs of possible worlds. This allows the unification of the various axioms. Let *w** denote the actual world; then we have the following fundamental translation schemata for PWS:

$\Box p$ → *p* is true at <u>every</u> *w* such that *w*Rw*

$\Diamond p$ → *p* is true at <u>some</u> *w* such that *w*Rw*

PWS allows of a reduction of all disputes concerning the modal axioms to disputes about the properties of the accessibility relations:

Old disputes give way to new. Instead of asking the baffling question whether whatever is actual is necessarily possible, we could try asking: is the relation R *symmetric?*

(Lewis, 1986, p. 19)

Thus, PWS is a tool that unifies and extensionalizes modal discourse and the diverse modal logics, i.e. it provides a unified extensional framework for talking about modal claims.

Correspondence Theory

Correspondence theory establishes 'relational translations' between the axioms of modal logic and certain properties of *R*. Of interest for us are the following correspondences:

Name	**Axiom** (modal formula)	**Property of *R***	**Characterisation**
T	$\Box p \rightarrow p$	Reflexive	$\forall w \in W : wRw$
D	$\Box p \rightarrow \Diamond p$	Serial	$\forall w \, \exists w_1 \in W : wRw_1$
B	$p \rightarrow \Box\Diamond p$	Symmetric	$\forall w_1, w_2 \in W : w_1Rw_2 \rightarrow w_2Rw_1$
4	$\Box p \rightarrow \Box\Box p$	Transitive	$\forall w_1, w_2, w_3 \in W : (w_1Rw_2 \,\&\, w_2Rw_3) \rightarrow w_1Rw_3$
5	$\Diamond p \rightarrow \Box\Diamond p$	Euclidean	$\forall w_1, w_2, w_3 \in W : (w_1Rw_2 \,\&\, w_1Rw_3) \rightarrow w_2Rw_3$

Reflexivity

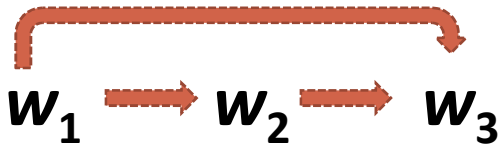
T is sound and complete in all the Kripke models in which the accessibility relation is **reflexive**.

Intuitively: all possible worlds are accessible from themselves.

Transitivity

S4 is sound and complete in the models whose accessibility relations are **reflexive** and **transitive.**

Transitivity guarantees that any world to which there is a patch from *w* is also directly accessible from *w*; truth at every world adjacent to *w* is enough to guarantee truth at every world to which there is a path from *w*.



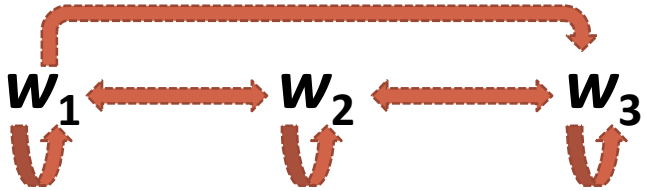
Symmetry

B is sound and complete in the models whose accessibility relation is **symmetric**.

Symmetry makes it that if *p* is true in a world *w*, then in every world w_n accessible from *w*, there is a world accessible from w_n in which *p* is true.

Equivalence

S5 is sound and complete in the models in which the accessibility relations are **equivalent**, i.e. reflexive, transitive, and symmetric.



Equivalence makes it that the class of worlds is split up into classes within which every world is accessible from every other world, there being no access between the classes

(*Wikipedia*)

Epistemic logic

The basic modal operator is ***K***, "it is known that," "it is epistemically necessary that"

$K_a p$ is read "agent *a* knows that *p,*" and $\neg K_a \neg p$ *is* read *"a* holds *p* possible" (= "*a* does not know that not *p*"):

Kp	::	$\Box p$
$\neg K \neg p$	::	$\Diamond p$

In a modal logic **<u>combining</u>** knowledge and belief, the 'translation' is as follows:

K	::	$\Box$
B	::	$\Diamond$

However, <u>in order to obtain a logic of belief, we first have to eliminate reflexivity as an accessibility relation *R*:</u> axiom T is dropped from **S5**!, and we have system **K45**. Adding axiom D to **K45**, we obtain system **KD45**.

Knowledge & Belief

Let us pretend that the notions of 'knowledge' and 'belief' are not problematic and let us thus define them under the perspective of epistemic modal logic:

Knowledge is veridical: $\boldsymbol{Kp} \rightarrow \boldsymbol{p}$ (axiom of knowledge; axiom *T*).

Belief is a propositional attitude of holding *p* as true.

Basically, $\boldsymbol{Kp} \rightarrow \boldsymbol{Bp}$, i.e. knowledge implies belief.

‘Strength’

Above (slide # 5), we spoke of strength of modal logic systems. In epistemic logic, we speak of ‘strength’ in the sense that the more properties of knowledge a system is capable of expressing via its axioms, the ‘stronger’ it is. Thus, while **K** is only capable of expressing the ***distribution property*** of knowledge (if one knows *p*, and knows that if *p* then *q*, then one also knows *q*), **T** is ‘stronger’ in that it expresses the ***knowledge property***, i.e. if one knows *p*, then *p* is true. Other properties of knowledge are the ***Brouwerian property*** expressed by axiom B, the ***consistency property*** expressed by axiom D, the ***positive introspection property*** (axiom 4), and, finally, the ***negative introspection property*** (axiom 5). According to this view, it is clear that systems **S4** and, especially, **S5** are considered the strongest; we can say that whatever is expressible in them has the most epistemic status.

Unconscious Knowledge & *t*-Accessibility

Clearly, if we are to speak of unconscious knowledge, i.e. the subject's lack of metaknowledge/ introspection, then the following axioms are not acceptable:

- **(4)** $Kp \rightarrow KKp$
- **(5)** $\neg Kp \rightarrow K\neg Kp$

That is to say that systems **S5** and **S4** are not capable of accounting for/incorporating unconscious knowledge. In fact, we already expected this, since we know that the metaknowledge of the subject increases with the more accessibility relations allowed. An 'ignorant knower' clearly has no access to his/her knowledge state, meaning that there is no **transitive** accessibility relation between possible worlds in the case of lack of metaknowledge.

t-Accessibility = introspection

This is so because **transitivity** describes **introspection:** in **S4** the subject has knowledge that s/he has knowledge (axiom 4), while in **S5**, when the subject does not know *p*, s/he knows that s/he does not know *p* (axiom 5). [Axiom 5 is Euclidean, but a relation that is reflexive and Euclidean is also symmetric and transitive, i.e. equivalent]

Unconscious Knowledge & *r*-Accessibility

If an 'ignorant knower' has no access to his/her knowledge state, then it seems intuitively reasonable to deny her/him also—or especially!—the reflexive accessibility relation: after all, we know that the subject lacks metaknowledge about his/her own epistemic state. However, axiom **T** holds *even* in the case of lack of metaknowledge:

(**T**) $\boldsymbol{Kp \rightarrow p}$

In fact, axiomatically, reflexivity simply means that <u>knowledge is veridical</u>. This simply eliminates belief.

Unconscious Knowledge & *s*-Accessibility

As for symmetry, in axiomatic terms it corresponds to the idea that if something is true, one knows one will not exclude it.

(B) $p \rightarrow K\neg K\neg p$

Clearly, 'ignorant knowers,' too, have access to this epistemic state concerning themselves.

D

But axiom D, $Kp \rightarrow Bp$, also holds—or should hold—in the case of lack of metaknowledge! This axiom expresses the epistemic situation that when one knows <that> *p*, then one believes <that> *p*, and we are actually committed to this claim **if we want to have a logic combining knowledge and belief**. However, what seriality might mean in epistemic terms is not intuitively clear: seriality as an epistemic accessibility relation means that for every world *w* where *p* is true, there's a world w_n where it is also true.

But, most importantly, sets expressing unconscious knowledge are not satisfiable in **D**!

A very ‘weak’ epistemic status…

A set expressing lack of metaknowledge would be, for instance, **{*Kp, ¬KKp*}**, only satisfiable in **T**.

The fact that sets expressing lack of metaknowledge are only satisfiable in system **T** and/or **K** is as good as attributing no epistemic status to unconscious knowledge, given that, as seen, these are the weakest epistemic systems; so weak, indeed, that for instance both {*p*, *Kp*, ¬*KKp*} and {*p*, *Kp*, *K*¬*Kp*} are consistent in **K** . Should this be a cause for concern in our stance that unconscious beliefs have positive epistemic status? Not really, because

Epistemic logic is no cause for concern 1

Scope and status of epistemic logic

- Epistemic logic describes the epistemic behaviour of merely *ideal* subjects (these always know what it is they know) (e.g.: Rescher, 1974);
- Despite worthy attempts (Levesque, 1984; Vardi, 1986; Fagin & Halpern, 1988), the problem in epistemic logic of logical omniscience is apparently unsolvable, thus emphasizing the *virtual* or *ideal* character of the epistemic behaviour described;
- "Referential opacity" and "the Logically Obtuse Man" invalidate every existing theorem of every epistemic logic (Hocutt, 1972);
- Epistemic logic is under fire concerning its status qua logic: "either [every theorem of epistemic logic] does not have anything especially to do with knowledge and is therefore epistemic in name only, or it does and is, in consequence, logic in name only" (Hocutt, 1972, p. 433);
- Knowledge implies beliefs *(Kp → Bp); however*, what we get in modal logic terms is two kinds of logic, epistemic and doxastic: the combination of both in a single modal logic is extremely problematic (Voorbraak, 1991).

Epistemic logic is no cause for concern 2

PWS and accessibility relations

- The epistemic accessibility relations are actually epistemically obscure, and not only at an intuitive level; in other words, “epistemic alternatives do not seem to be anything at all like possible worlds” (Konolige, 1986);
- PWS is far too problematic, namely in ontological terms, to provide a reliable tool of analysis of any kind;
- The concept of possible worlds expresses possibility and necessity; the translation of these into knowledge and belief is far from straightforwardly obvious, or justified.

Epistemic logic is no cause for concern 3

Correspondence Theory

- The problem of a correspondence algorithm is believed to be **undecidable** and it is **not arithmetical** (Szalas, 1994); this gives to the claimed 'correspondence' between modal schemata and first-order formulas a quite 'mysterious' character, to say the least;
- It is objected that **efficiency is lost** *because* of the translation: the structure of the original formulas, once translated, no longer holds (cf. Ohlbach, 1991, p. 692);
- Among other limitations, it does not appear to be able to provide a complete view of PWS, namely in what regards 'impossible' worlds (Pearce & Wansing, 1988).

Epistemic logic is no cause for concern 4

Last but not least:

Epistemically,

accessibility is not *availability* !

Even accepting it that epistemic logic can represent epistemic accessibility, namely as introspection,

it cannot represent availability of knowledge other than in terms of *logical inference*, and this gets us into one of the biggest problems faced by epistemic logic: **Logical Omniscience**!

References

- Augusto, L. (2008). "Do Unconscious Beliefs Yield Knowledge?" (forthcoming)
- van Benthem, J. (2006). "Epistemic Logic and Epistemology: The State of Their Affairs." *Philosophical Studies*, 128, p. 49-76
- de Bruin, B. (2007). "Epistemic Logic and Epistemology." In Hendricks, V. F. & Pritchard, D. (eds.), *New Waves in Epistemology*, Basingstoke: Palgrave Macmillan, p. 106-136
- Dienes, Z. & Perner, J. (2002). "A theory of the implicit nature of implicit learning." In French, R. M. & Cleeremans, A. (eds.), *Implicit Learning and Consciousness: An Empirical, Philosophical, and Computational Consensus in the Making?* New York: Psychology Press, p. 68-92
- Fagin, R. & Halpern, J. Y. (1988). "Belief, Awareness, and Limited Reasoning." *Artificial Intelligence* 34, p. 39-76
- Hendricks, V. F. & Symons, J. (2006). "Where's the Bridge? Epistemology and Epistemic Logic." *Philosophical Studies* 128, p. 137-167
- Hocutt, M. O. (1972). "Is Epistemic Logic Possible?" *Notre Dame Journal of Formal Logic* XIII:4, p. 433-453
- Konolige, K., (1986). "What Awareness Isn't: A Sentential View of Implicit and Explicit Belief." *Proceedings of the 1986 Conference on TARK, Monterey, CA,* San Francisco, CA: Morgan Kaufmann Publ., p. 241-250
- Levesque, H. J. (1984). "A Logic of Implicit and Explicit Belief." *Proceedings of the National Conference on Artificial Intelligence, Los Altos, CA*. Morgan Kaufmann Publ., p. 198-202
- Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Basil Blackwell.
- Ohlbach, H. J. (1991). "Semantics-Based Translation Methods for Modal Logics." *Journal of Logic and Computation* 1:5, p. 691-746
- Pearce, D. & Wansing, H. (1988). "On the Methodology of Possible Worlds Semantics, I: Correspondence Theory." *Notre Dame Journal of Formal Logic,* 29:4, p. 482-496
- Rescher, N. (1974). "On Alternatives in Epistemic Logic." *Studies in Modality*, *American Philosophical Quarterly*, monograph no. 8, Oxford: Basil Blackwell, p. 99-111
- Szalas, A. (1994). "On an automated translation of modal proof rules into formulas of the classical logic." *Journal of Applied Non-Classical Logics*, 4:2, p. 119-128
- Vardi, M. Y. (1986). "On Epistemic logic and Logical Omniscience." *Proceedings of the 1986 Conference on TARK, Monterey, CA,* San Francisco, CA: Morgan Kaufmann Publ., p. 293-305
- Voorbraak, F. (1991). "The Logic of Objective Knowledge and Rational Belief." In van Eijck, J. (ed.), *Logics in AI (Proceedings of JELIA '90)*, LNCS 478, Springer, p. 499-516