Epistemic logics for derived knowledge and belief

Julien Dutant*

Xixi Logic group, ZJU
Extended abstract
April 12th 2019

You have *derived* knowledge when you know something on the basis of something else. Inferential knowledge is a pardigm example of derived knowledge: if you come to know q by inference from $p_1, \ldots, p_n$, then you know q *on the basis of* $p_1, \ldots, p_n$. Derived knowledge raises for questions. (a) *Logical (non-)omniscience*: that q is entailed known $p_1, \ldots, p_n$ is not sufficient for knowing q on their basis. So what suffices? (b) *Inductive knowledge:* that q is entailed by $p_1, \ldots, p_n$ is not required for knowing q of their basis. So what is required? (c) *Closure*: a partial answer to (a) is that, when q is entailed known $p_1, \ldots, p_n$, by is additionally required is that one *competently deduces* q from $p_1, \ldots, p_n$. What is competent deduction? Is it sufficient for derived knowledge? (d) *Counter-closure*: a partial answer to (b) is that inference and basing cannot yield knowledge unless the premises or bases are known. In the deductive case, this is the *counter-closure* idea what is believes solely on the basis of premises that are not known is not known. Is knwoledge of the premises required?

The paper introduces a logic for derived knowledge to address these questions. The logic draws on the safety theory of knowledge (Sosa, 1999; Williamson, 2000) and significantly expands Williamson's "refined" models (Williamson, 2009). The core idea is to extend the notion of safety to attitudes to *arguments*. A belief is an attitude towards a proposition. It is *safe* just if its *epistemic counterparts* (beliefs that are like it in various epistemically relevant respects, such as how they

*King's College London. julien.dutant@kcl.ac.uk

are formed, which circumstances they are formed in, etc.) are true. Similarly, the 'argumentative' analogue of belief—which we may think of as a *conditional* belief (Edgington, 1995)—is a attitude towards *several* propositions: a conclusion and some premises. It is *safe* just if counterpart conditional beliefs are materially true: they have a true conclusion or some false premise. The second idea is to characterize *basing* in terms of *coordination of counterparts*. If you infer your belief that q from $p_1, \dots p_n$, then no belief is like it unless it is *similarly inferred from similar premisses*. That is: one bases one's belief that q on $p_1, \dots p_n$ only if the counterparts of one's belief that q are all conclusions in a countepart of one's attitude to the argument with counterparts of one's beliefs in the premises. Putting the two ideas together, we say that one *safely bases* one's belief that q on $p_1, \dots p_n$ just if one bases one's belief that q on $p_1, \dots p_n$ and one's attitude to the argument from q to $p_1, \dots p_n$ is safe. That, we claim, captures a notion of *knowledgeably basing* a belief on others: basing it in a such a way that if the premises are known, one can know the conclusion on their basis.

Formally, a frame for a language $\mathcal{L}$ is a pair $\langle W, R\rangle$ where W is a set of worlds and R a reflexive relation among *argumentative attitudes*, which are themselves triples $\langle w, q, \{p_1, \dots p_n\}\rangle$ of a world, a conclusion and a set of premises (we use formulas of $\mathcal{L}$ as propositions). This is enough to define three multiadic operators:

$S(q|p_1, \dots p_n)$ One's attitude to the argument from $p_1, \dots, p_n$ to q is safe.
: $w \models S(q|p_1, \dots, p_n)$ iff for all $w', q', p'_1, \dots p'_m$ such that $\langle w, q, \{p_1, \dots p_n\}\rangle R \langle w', q', \{p'_1, \dots p'_m\}\rangle$, $w' \models q$ or $w' \not\models p'_i$ for some $1 \leq i \leq m$.

$B(q|p_1, \dots, p_n)$ One bases one's belief in q on $p_1, \dots, p_n$.
: $w \models B(q|p_1, \dots, p_n)$ iff for all w', q' such that $\langle w, q, \varnothing\rangle R \langle w', q', \varnothing\rangle$, there are $p'_1, \dots p'_m$ such that $\langle w, q, \{p_1, \dots, p_n\}\rangle R \langle w', q', \{p'_1, \dots p'_m\}\rangle$ and $\langle w, \{p_1, \dots, p_n\}\rangle R^* \langle w', \{p'_1, \dots, p'_m\}\rangle$, where R^* captures the idea that the beliefs in $p'_1, \dots, p'_m$ at w' are like the beliefs in $p_1, \dots, p_n$ at w without enforcing one-to-one pairing.[1]

[1]Namely, R^* captures the "image" idea that every belief in $p_1, \dots, p_n$ at w has a counterpart in one's beliefs $p'_1, \dots, p'_m$ at w' and each of the latter is the counterpart of one of the former: for every $1 \leq i \leq n$ there is some $1 \leq j \leq m$ such that $\langle w, p_i, \varnothing\rangle R \langle w, p'_j, \varnothing\rangle$ and for every $1 \leq j \leq m$ there is some $1 \leq i \leq n$ such that $\langle w, p_i, \varnothing\rangle R \langle w, p'_j, \varnothing\rangle$.

$K(q|p_1, \ldots, p_n)$ One safely (knowledgeably) believes q on the basis of $p_1, \ldots, p_n$.[2]

$w \models K(q|p_1, \ldots p_n)$ iff *both* the condition for S and B above are satisfied.

We think of belief *simpliciter* as an attitude to an argument with an empty set of premises: $\langle w, q, \varnothing \rangle$. Hence monadic operators are simply variadic ones with no premises: Kp abbreviates $K(p|)$. Focusing on the K part of the fragment, we show that the following logic is sound and complete:

> An axiomatization of PL with *modus ponens* and the schemas:
> **MT**. $K(q|p_1, \ldots, p_n) \to (p_1 \wedge \ldots \wedge p_n \to q)$.
> **NEC.** $K(q|p_1, \ldots, p_n) \to (Kp_1 \wedge \ldots \wedge Kp_n \to Kq)$.[3]

Logics for the interaction of operators as well as stronger logics for subclasses of models will also be presented.

A picture of derived knowledge arises from the models. Here are some of its features.

The **MT** (material truth) axiom generalizes axiom **T** to argumentative attitudes.

Logical omniscience is entirely avoided: $K(p \to p)$ is not a theorem, nor $K(p|p \wedge p)$, for instance. Just like on good versions of safety, one can have an unsafe belief in a necessary truths, one can have an unsafe attitude in a logically valid argument.

NEC captures a closure idea: safe belief is closed under safe basing. The result is not trivial, as "safe basing" is not defined in terms of having a safe belief on some basis.[4] Thus safe basing is a good candidate for spelling out the notion of "competent deduction".

However, safe basing is not restricted to *deductive* arguments. From the point of view of the models, what matters for epistemic purposes is whether one's *attitude* to an argument or proposition is safe (whether its counterparts are true or materially true), not whether the argument is valid or the proposition logically true. The two are in principle orthogonal.

[2] We treat premises as sets: $B(q|p_1, p_1, p_2)$ is the same formula as $B(q|p_1, p_2)$. There is no restriction on embeddings: $K(K(q|p)|B(p|p_1, p_2))$ is a wff.

[3] When $n = 0$ treat $p_1 \wedge \ldots \wedge p_n$ and $Kp_1 \wedge \ldots \wedge Kp_n$ as being a tautology. Hence for $n = 0$ **NEC** is simply $Kq \to Kq$ and **MT** is the T axiom $Kq \to q$.

[4] In the models, $K(q|p)$ is compatible with $\neg Kq$. What is inconsistent is $K(q|p)$, Kp and $\neg Kq$.

Safe basing is a knowledge-like notion, not a "justification"-like notion. Hence we expect cases where one's inference is reasonable, materially true or even logically valid, and yet one fails to know because one is close enough to a mistake. We argue that this is a good diagnosis of some apparent counter-examples to closure (Lasonen-Aarnio, 2008).

The idea of a belief being *solely based* on a given argument from others can be expressed in the models too. But even if we add it the models fail to validate a *counter-closure* principle. Hence they provide an independent reason to be suspicious of counter-closure (comp. Warfield, 2005; Luzzi, 2010; Fitelson, 2010; Hawthorne and Rabinowicz, ming, a.o.).

The new logic thus promises to shed some light on how knowledge interacts with inference and reasoning.

References

Edgington, D. (1995). On conditionals. *Mind*, 104(414):235–329.

Fitelson, B. (2010). Strengthening the case for knowledge from falsehood. *Analysis*, 70(4):666–669.

Hawthorne, J. and Rabinowicz, D. (forthcoming). Knowledge and false belief. In de Almeida, R. B. C. and Klein, P., editors, *Explaining Knowledge: New Essays on the Gettier Problem*. Oxford University Press.

Lasonen-Aarnio, M. (2008). Single premise deduction and risk. *Philosophical Studies*, 141(2):157–173.

Luzzi, F. (2010). Counter-closure. *Australasian Journal of Philosophy*, 88:673–83.

Sosa, E. (1999). How to defeat opposition to moore. *Philosophical Perspectives*, 13:141–53.

Warfield, T. A. (2005). Knowledge from falsehood. *Philosophical Perspectives*, 19(1):405–416.

Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press.

Williamson, T. (2009). Probability and danger. *The Amherst Lecture in Philosophy*, pages 1–35.