Some Models for Safety and Closure

Julien Dutant
(University of Geneva)

Abstract

Drawing on Williamson (2009), we discuss some formal representations of safety conditions on knowledge that implement a popular pair of views, namely that safety is closed under competent deduction but not under logical entailment.

Many epistemologists take knowledge to require safe belief (Williamson, 2000; Sosa, 1999; Pritchard, 2005). Fewer take safe belief to be sufficient for knowledge. Here we sidestep the question whether safe belief is knowledge and we focus on spelling out safety itself.

Safety theorists want safety to be preserved by competent deduction: beliefs competently deduced from safe belief are themselves safe (Williamson, 2009, Sosa, 1999, 149). But they do not want safety to be closed under logical entailment: not anything that logically follows from what is safely believed is itself safely believed. Williamson (2009) proposes a formal semantics with these features. Williamson’s models have a puzzling feature: they require globally coordinated epistemic counterpart-hood relations. We explore a couple of alternatives that drop the requirement.

1 Safety

Williamson (2000, 100) states safety as follows:

(1) One’s belief that p in a case α is safe iff one avoids false belief cases sufficiently similar to α.

Williamson (2000, 94) takes cases to be centred worlds. But that will not do. Suppose that at a centred world α one has a true belief that p and a false belief that q. Since similarity is reflexive, there is a case similar to α in which one does not avoid false belief — namely, α itself. So one’s belief that p is not safe. The result is absurd: we do not want to deem all our beliefs unsafe as soon as

At least when the “deduction” is synchronous, i.e. more a matter of support than reasoning (Williamson, 2009, 26).
one of them is false. To avoid it, we should not reject reflexivity, which ensures
that safe beliefs are true. Nor should we say that safely believing \( p \) only requires
that one avoids false belief that \( p \) at sufficiently similar cases: that makes beliefs
in necessary truths trivially safe, an undesirable result (Williamson, 2009, 23).
Rather, we should use a notion of similarity that holds between beliefs (that is, belief
episodes) themselves. Provided that one’s belief that \( q \) is dissimilar enough from one’s belief that \( p \), the fact that one is mistaken does not tell
against the other being safe.

The idea can be spelled out with a finer-grained notion of case. Let a case
be a pair of centred world and a proposition. A case is a case of belief if the
subject of the case believes the proposition of the case at the time and world of
the case; it a case of truth if the proposition of the case is true at the centred
world of the case. We state safety:

Safety A case of belief is safe iff all relevantly similar cases are cases of truth.

You may have preferred: “... iff all relevantly similar cases of belief are cases of
truth.” If so, let it be a constraint on similarity that only cases of belief are similar
to cases of belief. Following Williamson (2009, 24), we can call relevantly similar
cases the “epistemic counterparts” of a case.

2 Williamson’s semantics

Williamson (2009, 24) formalizes Safety as follows. We have a language for
propositional logic supplemented with the \( K \) operator. We read \( Kp \) as “one’s
belief that \( p \) is safe”. A model is a triple \( (W, R, V) \) where \( W \) is a set of centred
worlds (henceforth “worlds”), \( V \) a valuation of the atoms, and \( R \) an “epistemic
counterpart” relation. The “epistemic counterpart” relation does not hold
between worlds as in standard Hintikka–Kripke semantics. Rather, it relates a
world to a world and a mapping of each sentence of the language to a sentence
of the language.

Let us explain. Suppose that \( R \) relates \( w \) to \( w^* \) and \( f \), where \( f \) is such a
mapping. \( f(p) \) may be \( p \) itself or some other sentence \( p \). The idea, then, is this:
one stands to \( f(p) \) in \( w^* \) relevantly like one stands to \( p \) in \( w \). In particular, if \( f(p) \)
is false in \( w^* \), then one’s standing with respect to \( p \) in \( w \) is too close to error.
The idea is reflected in the semantic clause for \( K \):

In the context of Williamson’s discussion, it is assumed that safe belief is
knowledge, but we leave that aside here.
\[ w \models Kp \text{ iff for all } (w, w', f) \in R, w' \models f(p). \]

(The rest of the semantics is as usual.) Reflexivity is obtained by saying that for every \( w, (w, w, 1) \in R \), where 1 is the function mapping every sentence to itself.

One may wonder where belief comes in. For all we have been told, \( Kp \) may hold at \( w \) without the subject of \( w \) even believing that \( p \) in \( w \). The idea can be captured within the confines of the model, though (see Williamson, 2009, 29 on “unconsidered” sentences). When you lack belief that \( p \), we treat your standing towards \( p \) is undifferentiated. Let \( 1_{p \rightarrow q} \) be the mapping that maps every sentence to itself except \( p \) which is mapped to \( q \). When you do not believe that \( p \) at \( w \), we say that \( (w, w, 1_{p \rightarrow q}) \in R \) for every \( q \) whatsoever. Since some sentence \( q \) is false at \( w \), you do not safely believe \( p \); \( w \models Kp \).

Williamson (2000, 31) establishes that the class of models have a sound and complete axiomatization, namely any set of axioms for propositional logic with modus ponens supplemented with the factivity schema \( Kp \rightarrow p \). Closure under entailment is avoided. Consider for instance a model with \( W = \{w\} \), \( V(p) = \{w\} \), \( R = \{(w, w, 1), (w, w, 1_{p \rightarrow \neg q})\} \): we have \( Kp \& p \) but not \( Kp \) (Williamson, 2009, 24).

Williamson defines a notion of safe derivation which he describes as an “analogue of competent deduction” (Williamson, 2009, 26). The relation obtains between sentences at a world:

Safe derivation \( q \) safely derives from \( p_1, \ldots, p_n \) at \( w \) iff whenever \( (w, w', f) \in R \) for some \( w \) and \( f \), if \( w \models f(p_1), \ldots, w \models f(p_n) \), then \( w \models f(q) \).

In intuitive terms, the definition says this. At \( w \), your standing to \( q \) safely derives from your standing to \( p_1, \ldots, p_n \) when the former and the latter are so coordinated that: when the counterparts of the latter are all true, the corresponding counterpart of the former is true.

Safe derivation preserves safety. It also has structural properties of a relation of logical consequence: Cut, Reflexivity and Monotonicity (Williamson, 2009, 28).

With one glitch: it allows truth-value gaps. Let \( (W, R, V) \) be such that \( W = \{w\} \), \( V(p) = \{w\} \), and \( R = \{(w, w, 1), (w, w, f)\} \) where \( f(p) = Kp \). The semantics tell us that \( w \models Kp \) if and only if \( w \models p \) and \( w \models Kp \). Since \( w \models p \), that boils down to: \( w \models Kp \) if and only if \( w \models Kp \); the semantics does not settle where \( Kp \) holds at \( w \). We can recast the semantics as a set of constraints on acceptable interpretations for the model; we say that a sentence holds at a world iff it holds on all acceptable interpretations.

Since \( (w, w, 1) \) is already included in \( R \), including \( (w, w, 1_{p \rightarrow q}) \) in \( R \) does not add new counterparts to sentences other than \( p \). So the inclusion does not affect the safety of anything else than \( (w, p) \).
Safe derivation relates to competent deduction as follows (Williamson, 2009, 28–9). Say that a set of sentences is logically transparent to one at w iff: all epistemic counterpart-hood relations at w preserve the logical relations of these sentences. For instance, if p, q and p&q are logically transparent to one at w, then for any f such that (w, w∗, f) ∈ R for some w∗, f(p&q) = f(p)&f(q). We have the following result:

If q, p₁,...,pₙ are logically transparent at w and p₁,...,pₙ entail q, then q safely derives from p₁,...,pₙ at w.

Insofar as competent deduction requires logical transparency, the result shows that safety is closed under competent deduction.

3 Two classes of weaker models

While elegant, Williamson’s models have a puzzling feature, namely that epistemic counterpart-hood is globally coordinated. Suppose that one’s belief that p at w is relevantly similar to some belief in p’ at w’. Williamson’s models require all propositions to have a counterpart at w. Moreover, they require that (w, p) is a counterpart to (w’, p) under (at least) one counterpart-hood relation such that any proposition at w has a unique counterpart at w’ under that relation.

Global coordination is not a superficial feature of the models. It plays a crucial role in the definition of safe derivation. Safe derivation says that when all the coordinated counterparts of the premises are true, the corresponding counterpart of the conclusion is true. Without global coordination, talk of “coordinated counterparts” and “corresponding counterpart” is not defined.

Prima facie, global coordination is puzzling. One may have a belief that p that has a counterpart in w and a belief that q that is simply without counterpart in w’. Can we do without it?

3.1 Uncordinated models

Looking back at Safety, a natural suggestion is to have an uncoordinated counter-part-hood relation between cases themselves. We use models (W, R, V) where W and V are as before, but R is reflexive relation between world–sentence pairs. The clause for safety is:

\[ w \models Kp \text{ iff for all } w, p \text{ such that } (w, p) \in R(w, p), w \models p. \]
The models turn out to be inter-translatable with Williamson’s, so they have the same logic. In uncoordinated models, we can define a notion of deductive inference akin to logical transparency.

Deductive inference \((u) q\) is deduced from \(p_1, \ldots, p_n\) at \(w\) iff: for any \(w', q'\)

\[
\text{such that } (w, q) \mathcal{R}(w', q'), \text{ there are } p_1, \ldots, p \text{ such that } (w, p_1) \mathcal{R}(w, p), \ldots \]

\[
\ldots (w, p_n) \mathcal{R}(w, p) \text{ and } p_1, p \text{ entail } q.
\]

In intuitive terms: \(q\) is deduced from \(p_1, \ldots, p_n\) at \(w\) iff every counterpart of \(q\) at \(w\) is a proposition that follows from some counterparts of \(p_1, \ldots, p_n\). It is easy to see that deductive inference preserves safety.

3.2 Locally coordinated models

Uncoordinated models do not give us a more general notion of inference. They allow us to state a plausible necessary condition on inference:

(\*) If at \(w\), one infers \(q\) from \(p_1, \ldots, p_n\), then for any \(w', q'\)

\[
\text{such that } (w, q) \mathcal{R}(w', q'), \text{ there are } p_1, \ldots, p \text{ such that } (w, p_1) \mathcal{R}(w, p), \ldots \]

\[
\ldots (w, p_n) \mathcal{R}(w, p) \text{ and } p_1, p \text{ entail } q.
\]

That is: there are counterparts of a conclusion belief only where there are counterparts of a premise belief. But the condition is not sufficient. Your belief that \(p\) may happen to have a counterpart wherever your belief that \(q\) has one without you having inferred \(q\) from \(p\).

The problem here is that inference does require some coordination. Suppose you inferred \(q\) from \(p_1, p_2\). That does not merely imply that there are counterparts of \(p_1, p_2\) wherever there are counterparts of \(q\). That also implies that the counterparts of \(q\) are related to some counterparts of \(p_1, p_2\) similarly to the way \(q\) is related to \(p_1, p_2\). In other words, we should look at counterparts of the pattern constituted by these three beliefs, not simply at individual counterparts of the beliefs.

The idea is captured by \textit{locally coordinated} models. These are like
Williamson’s models except that we allow partially defined mappings of sentences. The clause for safety is:

Here is the translation procedure. Given a globally coordinated model \((W, R, V)\), we get a pointwise equivalent uncoordinated model \((W, R, V)\) with \(R\) such that \((w, p)R(w, p')\) iff there is some \(f\) such that \((w, w', f)R^*\) and \(f(p) = p\). Given an uncoordinated model \((W, R, V)\), we get a pointwise equivalent globally coordinated model \((W, R^*, V)\) with \(R^*\) such that \((w, w', f)\in R^*\) iff \((w, p)R(w, p')\) and \(f\) maps \(p\) to \(p\) and every other sentence to some tautology.
\( w \models \text{Kp} \iff \text{for all } (w, w^* , f) \in R \text{ such that } f(p) \text{ is defined, } w^* \models f(p). \)

Reflexivity is given by ensuring that for every \( R, w, p \) there is a \( f \) such that \( (w, w, 1) \in R \) and \( f(p) = p \). Globally coordinated models are a special case where all mappings are required to be complete. Uncoordinated are a special case where all \( f \) are required to map only one sentence. The class of locally coordinated models has the same logic as the other ones.

We get a reasonable definition of inference:

Inference At \( w \), one infers \( q \) from \( p_1, \ldots, p_n \) iff for every \( w', f \) such that \( (w, w', f) \in R \) and \( f(q) \) is defined, \( f(p_1), \ldots, f(p_n) \) is defined.

That is, \( q \) is inferred from \( p_1, p_2 \) if and only if every counterpart of \( q \) is part of a counterpart of the "pattern" of beliefs (cases) \( q, p_1, p_2 \). No such definition is available on globally coordinated models.

An inference is safe if it preserves truth at its counterpart patterns:

Safe inference One safely infers \( q \) from \( p_1, \ldots, p_n \) at \( w \) iff one infers \( q \) from \( p_1, \ldots, p_n \) at \( t \), and for every \( w, f \) such that \( (w, w, f) \in R \) and \( f(q) \) is defined: if \( f(p_1), \ldots, f(p_n) \) are true at \( w' \), \( f(q) \) is true at \( w' \).

Safe inference preserves safety. Finally, we define deductive inference:

Deductive inference (l) At \( w \), one deduced \( q \) from \( p_1, \ldots, p_n \) iff for every \( w', f \) such that \( (w, w', f) \in R \) and \( f(q) \) is defined, \( f(p_1), \ldots, f(p_n) \) is defined and \( f(p_1), \ldots, f(p_n) \) entail \( f(q) \).

It is easy to see that Deductive inference is a Safe inference.

Cut and Assumption hold for our three notions of inference, but not Monotonicity. Suppose you inferred \( q \) from \( p \). To ensure Monotonicity for Inference, we need to require that for every \( w, f \) such that \( (w, w, f) \in R \) and \( f(q) \) is defined, \( f(r) \) is defined for any sentence \( r \). That would in effect turn our models in globally coordinated ones.

But the failure of monotonicity is intuitive. When you competently deduced \( q \) from \( p_1 \) and \( p_2 \), it does not thereby follow that you competently deduced \( q \) from \( p_1, p_2, r \) where \( r \) is some arbitrary proposition. Monotonicity is suitable for consequence, not inference.

For those we ensure Reflexivity by including every \( (w, w, f) \) such that \( f(p) = p \) and \( f \) is undefined otherwise. Using \( (w, w, 1) \) to ensure Reflexivity in our models ensures that safe inference (to be defined below) preserves truth and that deductive inference (to be defined below) requires entailment.
4 Conclusion

Locally coordinated models share the advantages of Williamson’s globally coordinated ones but afford a more perspicuous representation of inference and competent deduction.

References


