Logic, Language, and Computation 2014

Close Encounters with Non-Normal Worlds

Franz Berto

F.Berto@uva.nl **f.berto@abdn.ac.uk**

0. What are they? (Definitions)

1. Why should we have them around?

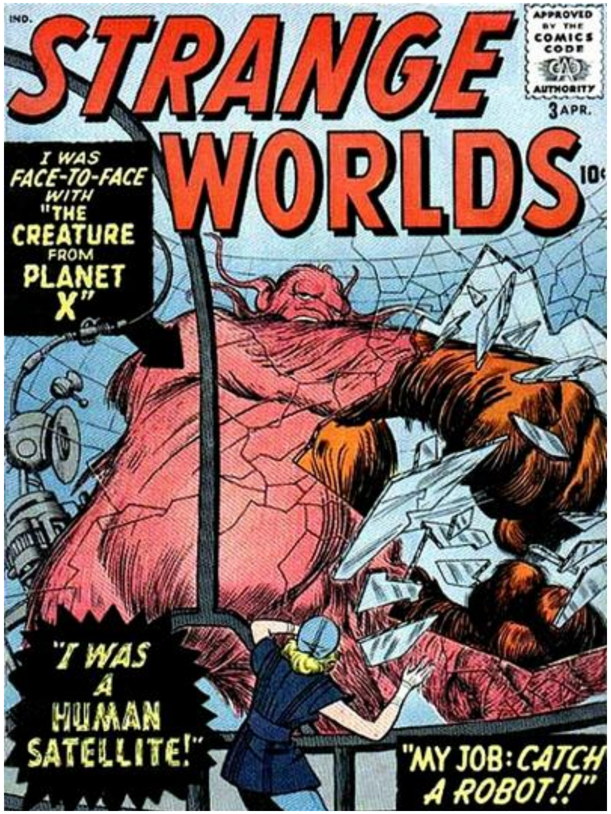
2. The logic(s)

3. Back to metaphysics

4. Further sample applications

5. Further readings

0. What are they?



1. **Ways things could not have been.** (Salmon 1984; Yagisawa 1988; Restall 1997; Beall & van Fraassen 2003)

2. **Worlds where (your favourite) logic fails**. Take some logic L: an impossible world with respect to the L-laws is one in which those laws fail to hold (Priest 2001)

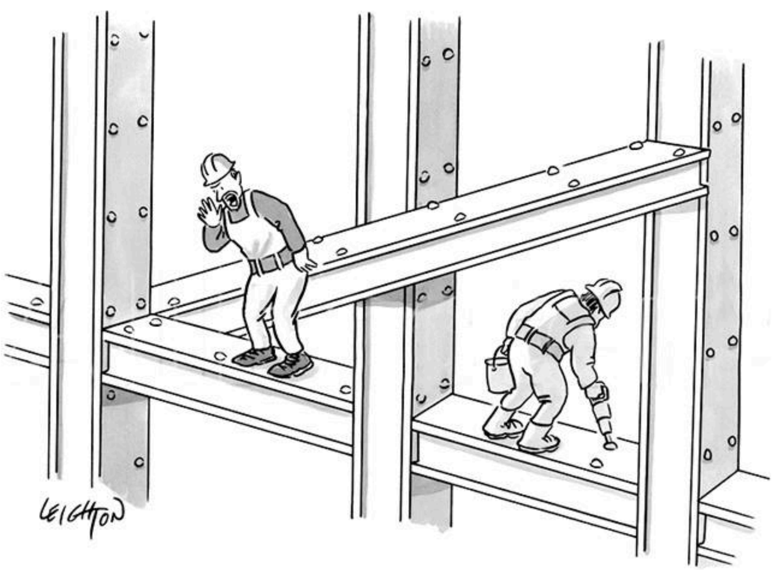
3. **Worlds where classical logic fails.** (Priest 1997a)

4. **Worlds making contradictions true**. (Rescher & Brandom 1980, Lycan 1994, Berto 2007)

Not any 1-impossible world is a 2-impossible world. Suppose the AC of set theory is true, and that mathematical necessity is unrestricted. Then a world where the AC fails may be 1-impossible but not 2-impossible.
An intuitionistic world can be 2- or 3- impossible without being 4-impossible.

(Yeah, but *what* are they?? What kind of stuff is an absolute impossibility supposed to be made of? – We'll get back to this.)

1. Why should we have them around?



"Escher! Get your ass up here."

- **The Argument from Ways** (Naylor 1986, Yagisawa 1988): “'ways' talk goes both ways” (Beall and van Fraassen 2003: 86): if quantification over ways the world could be (Lewis 1973) provides evidence for possible worlds, then quantification over ways the world could not be provides evidence for IWs.

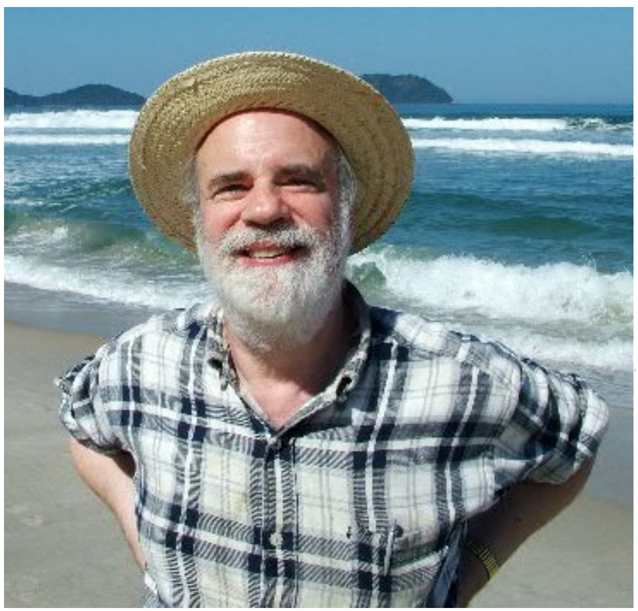
  Convincing?

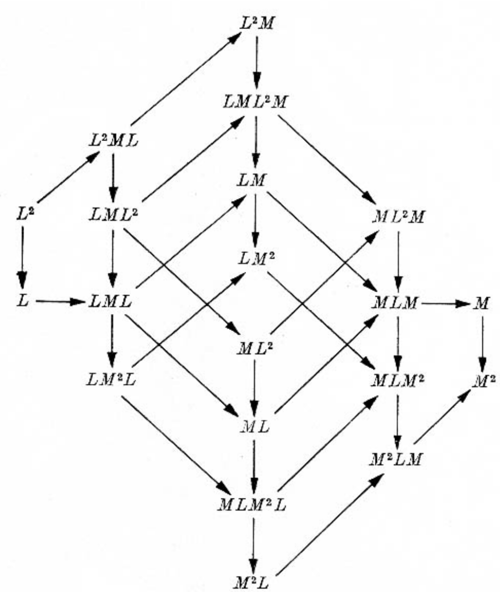
- **The Argument from Counterpossibles** (Nolan 1997, Brogaard & Salerno 2013): “If Hobbes had squared the circle, then all mathematicians would have been amazed” (true!); “If Hobbes had squared the circle, then Kennedy would not have been killed” (false!). You need IWs for this.

  Convincing?

- **The Argument From Utility**: you can do *a lot of stuff* using IWs. Let’s see…

2. The logic(s)




L^2M
LML^2M
L^2ML
LM
L^2
LML^2
ML^2M
LM^2
L
LML
MLM
M
ML^2
LM^2L
MLM^2
M^2
ML
M^2LM
MLM^2L
M^2L

2.1. IWs in non-normal modal logics

C.I. Lewis' S2, S3, and other non-normal modal systems do not include the Rule of Necessitation. Model−theoretically :

(NEC) If $\models A$, then $\models \Box A$.

But even in K, this holds…

IWs to the rescue!

Kripke 1965 introduced non-normal worlds to deal with (NEC).
Take a tuple $\langle W, N, R, v \rangle$, where
W is a set of worlds;
N is a proper subset of W, the set of normal worlds;
R is a binary accessibility;
v is a valuation function: "$v_w(A)$" = the truth value of A at w.
Worlds in W − N are the non−normal folks.

2.1. IWs in non-normal modal logics

□ and ◊ get the usual truth conditions at normal worlds. But at non−normal worlds formulas of the form £A, with £ a modal, are not evaluated recursively depending on the truth value of A at other (accessible) worlds, but get assigned their truth value *directly*.

Specifically, if w is a non−normal world, the truth conditions for the modalizers are :

$v_w(\Box A) = 0$
$v_w(\Diamond A) = 1$

Where 1 stands for true, 0 for false. So all box−formulas are false and all diamond−formulas are true: these are worlds where nothing is necessary, but anything is possible.

2.1. IWs in non-normal modal logics

Logical validity / consequence are holding / truth preservation at *normal* worlds, thus (with S a set of formulas):

$\vDash$ A iff, for all interpretations $\langle W, N, R, v\rangle$, and all worlds $w \in N$, $v_w(A) = 1$.

S $\vDash$ A iff, for all interpretations $\langle W, N, R, v\rangle$, and all worlds $w \in N$, if $v_w(B) = 1$ for all formulas $B \in S$, then $v_w(A) = 1$.

(We want to look only at possible or normal worlds, that is, worlds where logic is *not* different, when we define what holds according to logic!)

(NEC) fails: take any classical tautology, say $A \vee \neg A$. This holds at all worlds of all interpretations, so $\vDash A \vee \neg A$. Therefore, $\Box(A \vee \neg A)$ holds at all normal worlds of any interpretation, so $\vDash \Box(A \vee \neg A)$. But $\Box(A \vee \neg A)$ does not hold in any non−normal world. Therefore, $\Box\Box(A \vee \neg A)$ is false at normal worlds that have access, via R, to any non−normal world, and so $\nvDash \Box\Box(A \vee \neg A)$.

2.2. IWs in epistemic logics

Because K (knowledge, or belief) is a restricted quantifier on *possible* worlds in standard Hintikka-epistemic logic, it suffers from logical omniscience.

(Closure) If KA, and $A \vDash B$, then KB

One knows (believes) all the logical consequences of the things one knows (believes). As a special case, all valid formulas are known (believed):

(Validity) If $\vDash A$, then $\vDash KA$

And beliefs form a consistent set (given Seriality), that is, it cannot be the case that both a formula and its negation are believed:

(Consistency) $\vDash \neg(KA \wedge K\neg A)$.

2.2. IWs in epistemic logics

This is *not* how real, finite, and fallible epistemic agents work!

- We all experience having (perhaps covert) inconsistent beliefs.
- Even though A ∨ ¬A, is (assume) logically valid, my intutionist friends do not believe it.
- We know such basic arithmetic truths as Peano's postulates; and Peano's postulates may entail (assume) Goldbach's conjecture; but we don't know whether Goldbach's conjecture is true.

IWs to the rescue!

2.2. IWs in epistemic logics

A *Rantala* (1982) *interpretation* is $\langle W, N, R, v \rangle$, where
W is our usual set of worlds;
N is the subset of normal, possible worlds;
W − N is the set of impossible worlds;
R is the accessibility.

Now the worlds accessible from a given w are worlds compatible with what the relevant cognitive agent believes at w, or with the evidence it has there, etc.

At IWs in W − N *all* formulas are assigned a truth value by v directly, not recursively: compound formulas of the form ¬A, A ∨ B, etc., behave arbitrarily: A ∨ B may turn out to be true even though both A and B are false (IWs may be nonprime), and ¬A may turn out to be true when A is (IWs may be inconsistent).

Logical consequence is, again, defined on *possible* worlds. IWs are logically completely anarchic.

2.2. IWs in epistemic logics

By allowing such worlds to be accessible via R in the evaluation of formulas including intentional−epistemic operators such as K, one can destroy their unwelcome closure features, thereby dispensing with Closure, Validity, and Consistency.

As for Consistency, for instance, just access via R an impossible world where both A and ¬A are true.

2.3. IWs in relevant logics

Take the infamous *fallacies of relevance*, or *paradoxes of the material and strict conditional*: these turn out to be true in classical logic just because the antecedent is a (necessary) falsity, and without any real connection between antecedent and consequent.

Famous is *ex contradictione quodlibet*, also called the Law of Explosion:

$A \wedge \neg A \rightarrow B$

Other irrelevant conditionals are those that turn out to be logical truths just because the consequent is necessary (*verum ex quolibet*), such as:

$A \rightarrow B \vee \neg B$
$A \rightarrow (B \rightarrow B)$

IWs to the rescue!

2.3. IWs in relevant logics

A Routley–Meyer interpretation (Routley & Routley 1972; Routley & Meyer 1973, 1976; Routley 1979) for relevant (propositional) logics is a quintuple ⟨W, N, R, *, v⟩, where
W is a set of worlds;
N is a proper subset of W including the normal or possible worlds;
W − N is the set of non–normal or impossible worlds;
R is a *ternary* accessibility relation defined on W;
* (the so–called *Routley star*) is a monadic operation on W, sometimes called involution.

2.3.1. Relevant conditional

IWs can be seen as scenarios where logical laws may fail, and the Law of (propositional) Identity $B \rightarrow B$ is one of them.

At *possible* worlds, we still require for the truth of conditionals $A \rightarrow B$ that at every accessible world where A holds, B holds, too. So $A \rightarrow (B \rightarrow B)$ is not logically valid.

That's the *insight*. Technically, it's a bit complicated.

When w is an impossible world, we state the truth conditions for the conditional, by means of the ternary R, thus:

(S→) $v_w(A \rightarrow B) = 1$ iff, for all worlds w_1 and $w_2 \in W$, such that Rww_1w_2, if $v_{w1}(A) = 1$, then $v_{w2}(B) = 1$.

2.3.1. Relevant conditional

B → B fails at w, when this is an impossible world such that for some worlds (which may be possible or not) w_1 and w_2, such that Rww_1w_2, B holds at the former and fails at the latter.

(What the hell does that ternary R *mean*? Look at *information flow*: Restall 2000, Mares 2004. Think of worlds as information states or data bases. When Rww_1w_2, w allows information to flow from w_1 to w_2. So if A holds at w_1, and w allows the information that A → B to flow from w_1 to w_2, then B should hold at w_2).

2.3.2. De Morgan negation

The truth conditions for negation are:

(S¬) $v_w(\neg A) = 1$ iff $v_{w^*}(A) = 0$

¬A is true at w if and only if A is false, not at w itself, but at its twin w*.

By assuming w** = w, one can validate Double Negation. The operator so characterized is often called *De Morgan negation*, for also De Morgan's Laws hold.

But it does not validate the Law of Explosion: consider a model in which A holds at w, B doesn't hold at w, and A doesn't hold at w*. Then, both A and ¬A hold at w, whereas B doesn't: w is an inconsistent but non−trivial world.

2.3.2. De Morgan negation

Now what does *this* * mean??

The twins are "mirror images one of the other reversing 'in' and 'out'" (Dunn 1986: 191). The reverse twin of a w which is A−inconsistent is world, w*, which is A−incomplete, and vice versa: involution takes local inconsistency into local incompleteness, and vice versa.

For some world w, it may also be the case that w = w*: the twins are in fact one. Then w just is a maximal consistent world. At such a w, negation behaves *completely classically*: ¬A is true at it if and only if A is false at *it*.

3. Back to metaphysics…

The Parity Thesis:

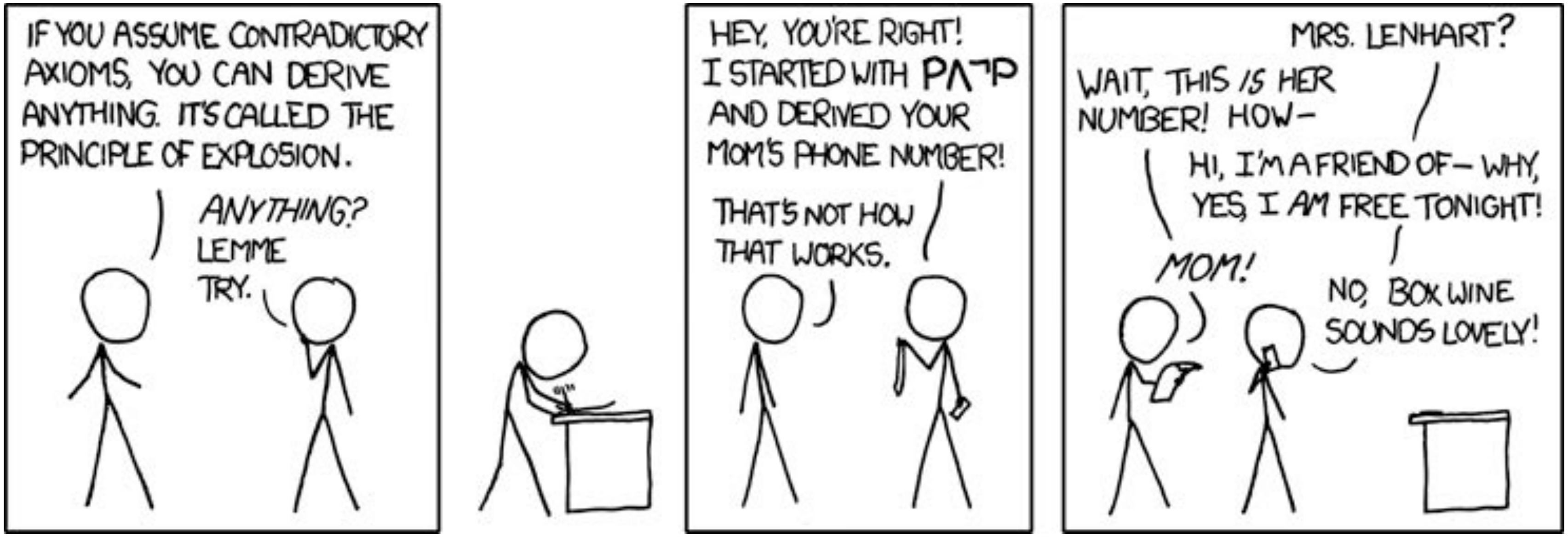
"As far as I can see, any of the main theories concerning the nature of possible worlds can be applied equally to impossible worlds: they are existent nonactual entities; they are nonexistent objects; they are constructions out of properties and other universals; they are just certain sets of sentences. ... There is, as far as I can see, absolutely no cogent (in particular, non−question−begging) reason to suppose that there is an ontological difference between merely possible and impossible worlds." (Priest 1997b: 580–1)

Yagisawa's *extended modal realism* proposes a Lewisian realist account of IWs and impossibilia: IWs, just like Lewis' possible worlds, are largely concrete maximal mereological sums of individuals, causally and spatiotemporally isolated from each other (Yagisawa 1988)

Impossibilist ersatzism: impossible worlds are ersatz constructions, abstract entities on a par with ersatz possible worlds (Mares 1997, Vander Laan 1997).

This option embeds various sub–options, for modal ersatzism comes in various shapes (Divers 2002, Part III): sets of propositions, of sentences from a "worldmaking" language, maximal non-obtainable states of affairs, etc.

4. Further sample applications



4.1. Propositional content

In possible worlds semantics propositions are functions from worlds to truth values, or sets of worlds: a proposition is the set of worlds at which it is true.

This has a notorious "granularity problem" (Barwise 1997) with impossible propositions: intuitively distinct impossible propositions (that swans are blue and it is not the case that swans are blue, that Fermat's Last Theorem is false, that Charles is a married bachelor) hold at the same possible worlds: none.

Of course, we have a dual problem with (unrestrictedly) necessary propositions, that are all identified as the total set of worlds.

… *IWs to the rescue!*

That the proposition expressed by A is impossible does not mean that it is an empty set of worlds, but rather that it includes only IWs. We can have an IW, w_1, with inconsistent swans; a distinct IW, w_2, at which Fermat's Last Theorem is false; and a still distinct IW, w_3, at which bachelors are married but swans and Diophantine equations behave wisely.

4.2. Counterpossible reasoning

The Lewis−Stalnaker theories of counterfactuals have it that "If it were the case that A, then it would be the case that B" is true iff at the closest world − or worlds − at which A is true, B is true (with Limit Assumption turned on, see Lewis 1973).

Then any counterfactual whose antecedent is impossible, that is, true at no possible world, is vacuously true: there being no worlds at which A is true, any closest A−world is trivially a B−world.

But we often need to nontrivially *reason* about theories which (perhaps unbeknownst to us) cannot possibly be correct, that is, to reason from antecedents that may turn out to be not only false, but necessarily so:

- Alternative logics;
- Mathematical conjectures;
- Metaphysical views.

4.2. Sample applications: Counterpossible reasoning

Much metaphysical talk is made with our quantifiers "wide open", that is, aiming at stating truths on *all* that there was, is, or could possibly be.

This is evident in modal ontology, when people advance a theory on the totality of worlds and on their nature. But other metaphysical debates easily come to mind.

If a philosopher is to evaluate metaphysical theories which she considers wrong (say, in order to draw unpalatable consequences), such as Spinoza's monism or Hegel's metaphysics of the Absolute, then she must envisage situations where such metaphysics are correct, and wonder what would be the case at them: situations at which there is only one substance, or at which the Absolute Geist necessarily shapes the teleological development of history…

4.2. Sample applications: Counterpossible reasoning

Semantic structures for counterfactual conditionals involving impossible worlds are in Routley 1989, Read 1995, Mares & Fuhrmann 1995, Mares 1997, Nolan 1997, Brogaard & Salerno 2013.

Most of these are natural extensions of Lewis' 1973 semantics for counterfactuals, and capture several intuitions on counterfactual conditionals with impossible antecedents and counterpossible reasoning.

The main task for such theories consists in accounting for the concepts of *closeness* and *similarity* between worlds once IWs enter the stage (Vander Laan 2004).

5. Further readings

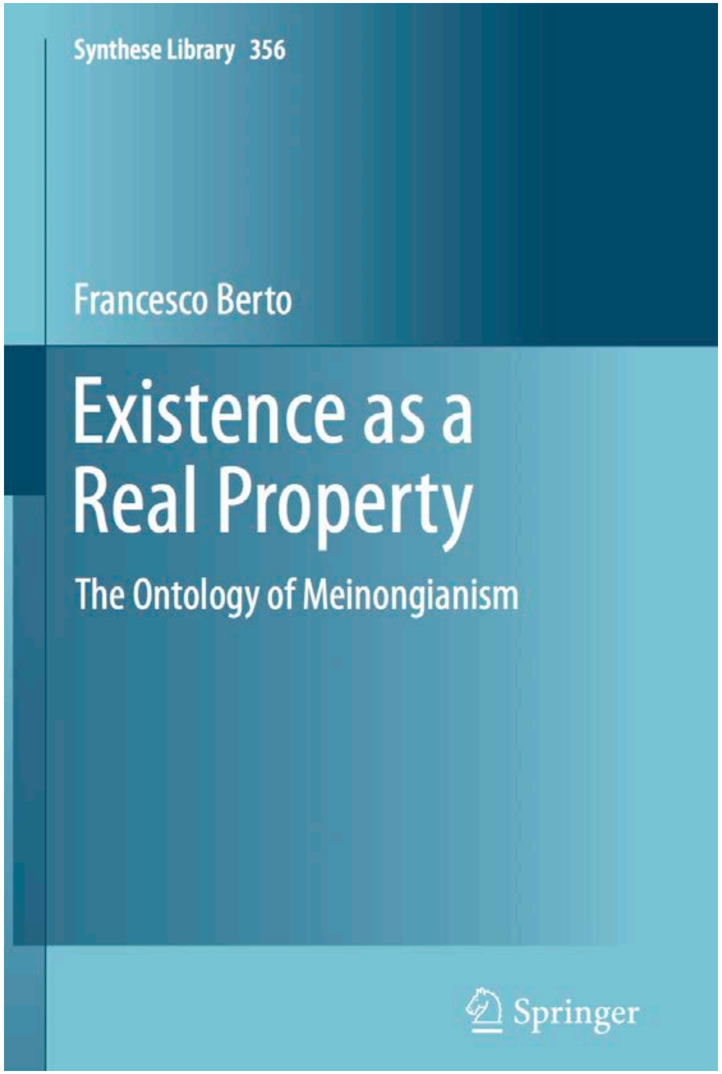
Extended intro with extended bibliography:

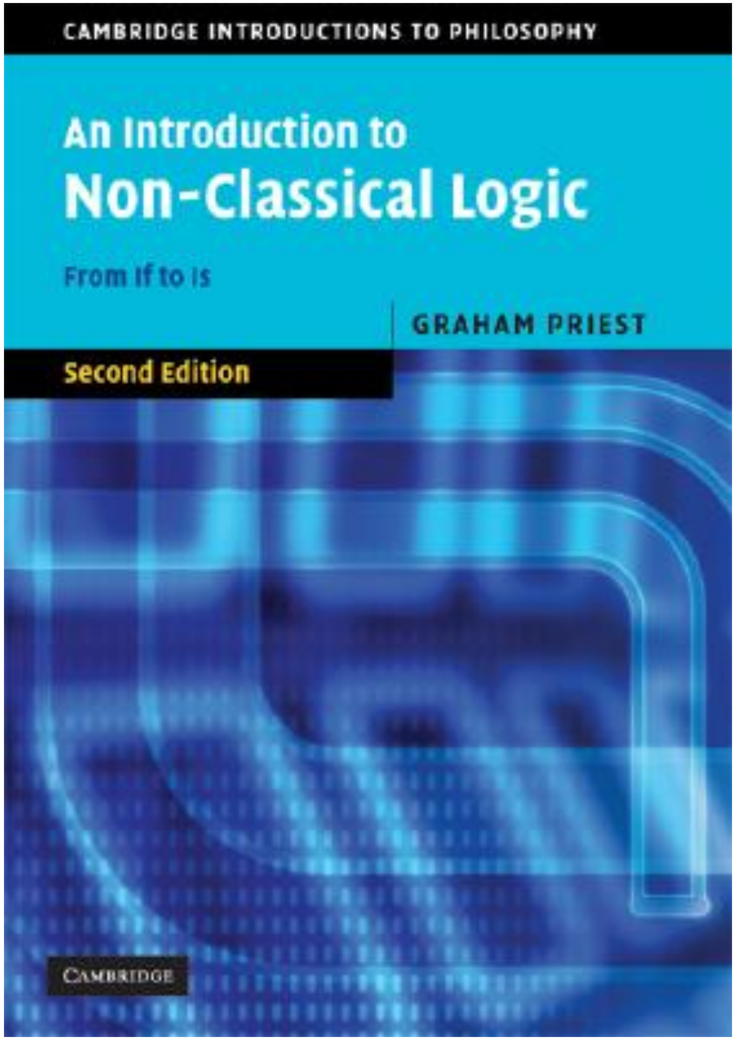
- "Impossible Worlds", *The Stanford Encyclopedia of Philosophy*, CSLI, Stanford, CA, 2013. http://plato.stanford.edu/entries/impossible-worlds

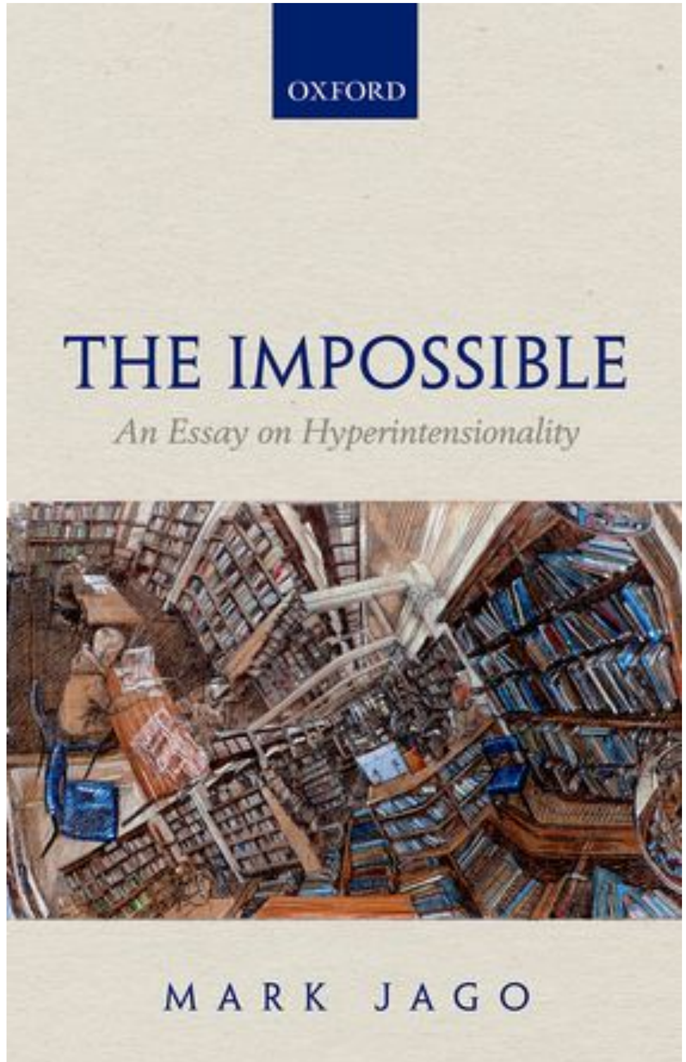
(Look at Section 6, on the *objections* to IWs!)

Further stuff from me:

- "A Modality Called 'Negation'", *Mind* (forthcoming 2015).
- "On Conceiving the Inconsistent", *Proceedings of the Aristotelian Society*, 114(2014): 103-21.
- "Non-Normal Worlds and Representation", *The Logica Yearbook 2011*, College Publications, London 2012: 15-30.
- "Impossible Worlds and Propositions: Against the Parity Thesis", *The Philosophical Quarterly*, 60(2010): 471-86.


Synthese Library 356
Francesco Berto
Existence as a Real Property
The Ontology of Meinongianism
Springer


CAMBRIDGE INTRODUCTIONS TO PHILOSOPHY
An Introduction to
Non-Classical Logic
From If to Is
GRAHAM PRIEST
Second Edition
CAMBRIDGE


OXFORD
THE IMPOSSIBLE
An Essay on Hyperintensionality
MARK JAGO