

*Forthcoming in the Australasian Journal of Logic*

# Logically Impossible Worlds

*Koji Tanaka*

School of Philosophy  
Research School of Social Sciences  
Australian National University  
Koji.Tanaka@anu.edu.au

---

## 1 Impossible Worlds

If we think of a possible world as representing a way things can be, an impossible world represents a way in which things cannot be. The importance of impossible worlds has been discussed in recent literature.<sup>1</sup> One reason for accepting impossible worlds comes from counterfactual reasoning (Nolan (1997), Routley (1989)). Consider, for instance, the following counterfactual: if someone were to square a circle, she or he would be famous. This counterfactual has an impossible antecedent. One way to analyse such a counterfactual is to invoke an impossible world where someone does square a circle and evaluate the consequent at the world or a set of worlds that is or are the closest to that impossible world.

But what exactly are impossible worlds? What does it mean to talk about ways in which things cannot be? There are several formulations of impossible worlds in the literature.<sup>2</sup> Some characterise the impossibility of impossible worlds in terms of logic. There have been mainly three ways in which impossible worlds are characterised in logical terms. First, an impossible world is a world where  $A \wedge \neg A$  is true for some  $A$  (Lycan (1994)). According to this view, an impossible world is a world where some contradiction obtains. Second, an impossible world is a world that is governed by some non-classical logic (Cresswell (1973)). According to this second formulation, it is assumed that classical logic holds at the actual world, and a world is impossible if the laws of classical logic fail. Because of the assumption that the actual world is governed by classical logic, this means that an impossible

---

<sup>1</sup>For a survey, see Berto (2013) §1.

<sup>2</sup>See Berto (2013), Jago (2015), Nolan (2013).

world is a world where the laws of logic are different. Third, an impossible world is a world where the laws of logic fail (Priest (2008): 172).<sup>3</sup>

The first and the second formulations of impossible worlds in terms of logic give us an insight into what an impossible world might be like. If impossibility is characterised in terms of contradictions, an impossible world is contradictory. It might be a world where Sylvan's Box, a box which is both empty and nonempty, exists (Priest (1997)). An impossible world in this sense might be a world which is governed by some paraconsistent logic, or classical logic if it is also a trivial world where everything is true. If an impossible world is a world which is governed, for instance, by intuitionistic logic, it is a world where mathematics is done intuitionistically. At such a world, a mathematician cannot rely on the Law of Excluded Middle to prove anything.<sup>4</sup> But what does it mean for the laws of logic to fail?

My task in this paper is to answer this question. I will use the resources that Routley/Sylvan developed with his collaborators for the semantics of relevant logics to explain a world where the laws of logic fail. I will claim that the *non-normal worlds* that Routley/Sylvan (with his collaborators) introduced are exactly such worlds.<sup>5</sup> To disambiguate different kinds of impossible worlds, I will call such worlds *logically impossible worlds*. At a logically impossible world, the laws of logic fail. In this paper, I will provide a definition of logically impossible worlds. I will then show that there is nothing strange about admitting such worlds.

## 2 Logically Impossible Worlds

We can define logically impossible worlds by elaborating on some features of the semantics for relevant logics that were introduced by Routley/Sylvan (with his collaborators). For simplicity, I consider the simplified semantics for the relevant logic *B* (Priest and Sylvan (1992)). The semantics for *B* is a structure  $\langle W, N, R, I \rangle$  where *W* is a set of worlds, *N* is a set of *normal* worlds,  $R \subset W \times W \times W$ , and *I* is a map from propositions to worlds. *R* at normal worlds satisfies the *normality condition*:

$$\langle w, w', w'' \rangle \in R \text{ iff } w' = w'' \text{ for every } w \in N \text{ and every } w', w'' \in W.$$

In other words, *R* at normal worlds is binary and at non-normal world is ternary.

---

<sup>3</sup>As we will see later, Priest (2008) equates a world where the laws of logic fail with a world where the laws of logic are different. In the last section of this paper, I will argue that that is a mistake.

<sup>4</sup>See Detlefsen (1990) for gaining an understanding of what that world might be like.

<sup>5</sup>Nolan (2013) makes a brief suggestion along this line (p. 361) but no sustained discussion has been given in the literature as far as I know.

For the sake of illustration, I will consider a so-called relational semantics and assume an evaluation of propositions to be not a function but a *relation* between a proposition and a value. So  $I$  maps a proposition,  $p$ , to an *extension*  $E_p$  (the set of worlds where  $p$  is true) and an *anti-extension*  $A_p$  (the set of worlds where  $p$  is false), i.e.,  $I(p) = \langle E_p, A_p \rangle$ . A two-valued semantics requires the Routley-star operation to specify the truth condition for negation. As will become clear, however, negation is not what is at issue in defining logically impossible worlds. In order to focus on the main issue, I will examine a relational semantics.<sup>6</sup>

An evaluation of a proposition,  $\rho$ , can be given as:

$$\begin{aligned} \langle p, 1 \rangle \in \rho_w &\text{ iff } w \in E_p. \\ \langle p, 0 \rangle \in \rho_w &\text{ iff } w \in A_p. \end{aligned}$$

Evaluations of complex formulas involving conjunction and disjunction also come in two clauses: one for extension and another for anti-extension.

Our main interest is with conditionals as an examination of conditionals can give us a definition of logically impossible worlds. Conditionals are evaluated in terms of ternary relations:

$$\begin{aligned} \langle A \rightarrow B, 1 \rangle \in \rho_w &\text{ iff for every } w', w'' \in W \text{ such that } \langle w, w', w'' \rangle \in R, \text{ if} \\ &\langle A, 1 \rangle \in \rho_{w'} \text{ then } \langle B, 1 \rangle \in \rho_{w''}. \\ \langle A \rightarrow B, 0 \rangle \in \rho_w &\text{ iff for some } w', w'' \in W \text{ such that } \langle w, w', w'' \rangle \in R, \\ &\langle A, 1 \rangle \in \rho_{w'} \text{ and } \langle B, 0 \rangle \in \rho_{w''}. \end{aligned}$$

Because of the normality condition, the evaluations of conditionals at normal worlds can be simplified as:

$$\begin{aligned} \langle A \rightarrow B, 1 \rangle \in \rho_w &\text{ iff for every } w' \in W, \text{ if } \langle A, 1 \rangle \in \rho_{w'} \text{ then } \langle B, 1 \rangle \in \rho_{w'}. \\ \langle A \rightarrow B, 0 \rangle \in \rho_w &\text{ iff for some } w' \in W, \langle A, 1 \rangle \in \rho_{w'} \text{ and } \langle B, 0 \rangle \in \rho_{w'}. \end{aligned}$$

Validity is then defined in terms of truth-preservation at all *normal* worlds:

$$\Sigma \models A \text{ iff for every } \langle W, N, R, I \rangle \text{ and every } w \in N, \text{ if } \langle B, 1 \rangle \in \rho_w \text{ for every } B \in \Sigma, \text{ then } \langle A, 1 \rangle \in \rho_w.$$

Note that because we have a relational semantics, falsity does not exclude the sentence being true as well. In order to show that a conditional, for instance  $A \rightarrow B$ , fails to be a logical truth at a world, we need to demonstrate not that it is false,  $\langle A \rightarrow B, 0 \rangle \in \rho_w$ , but that it fails to be true at that world, i.e.,  $\langle A \rightarrow B, 1 \rangle \notin \rho_w$ .

---

<sup>6</sup>The relevant logic based on a relational semantics is, in fact, not the same as  $B$ . The relational semantics generates a logic that does not validate contraposition whereas  $B$  does. For details, see Restall (1995).

Now, a conditional is said to be *relevant* if the antecedent and the consequent share some propositional variables. For a conditional to be a logical truth of relevant logics, it must be relevant. The following conditionals fail to be logical truths according to relevant logics:  $A \rightarrow (B \vee \neg B)$ ,  $(A \wedge \neg A) \rightarrow B$ , and  $A \rightarrow (B \rightarrow B)$ .

The first conditional ( $A \rightarrow (B \vee \neg B)$ ) can be invalidated by allowing some worlds to be incomplete: neither  $B$  nor  $\neg B$  are true. If  $A$  is true at such an incomplete world, the antecedent of the conditional ( $A$ ) is true but the consequent ( $B \vee \neg B$ ) is not true at that world. So, for some  $w' \in W$ ,  $\langle A, 1 \rangle \in \rho'_{w'}$  but  $\langle B \vee \neg B, 1 \rangle \notin \rho'_{w'}$ . Hence, for some normal world  $w \in N$ ,  $\langle A \rightarrow (B \vee \neg B), 1 \rangle \notin \rho_w$ . Once we introduce incomplete worlds to the set of worlds, the conditional can be shown to be invalid.

The second conditional ( $(A \wedge \neg A) \rightarrow B$ ) can be invalidated in terms of inconsistent worlds where both  $A$  and  $\neg A$  are true. If  $B$  is not true at such an inconsistent world, the conditional can be shown to be not true. Suppose that  $\langle A \wedge \neg A, 1 \rangle \in \rho_{w'}$  for some  $w' \in W$ . If  $\langle B, 1 \rangle \notin \rho_{w'}$ , then  $\langle A \wedge \neg A, 1 \rangle \in \rho_{w'}$  and  $\langle B, 1 \rangle \notin \rho_{w'}$  for some  $w' \in W$ . So, for some normal world  $w \in N$ ,  $\langle (A \wedge \neg A) \rightarrow B, 1 \rangle \notin \rho_w$ . So the conditional is invalid.

From the perspective of relevant logics, it is the third conditional,  $A \rightarrow (B \rightarrow B)$ , that is the most interesting. In most relevant logics,  $B \rightarrow B$  comes out to be a logical truth. For every  $w' \in W$ , if  $\langle B, 1 \rangle \in \rho_{w'}$  then  $\langle B, 1 \rangle \in \rho_{w'}$ . So  $\langle B \rightarrow B, 1 \rangle \in \rho_w$  for any  $w \in N$ . Hence the conditional is valid. So  $A \rightarrow (B \rightarrow B)$  has a consequent that is a logical truth. This is where the distinction between normal and non-normal worlds becomes important.

Routley and Routley (1972) developed semantics for *First Degree Entailment* (FDE). FDE is a conditional-free fragment of (full) relevant logics. When Routley and Meyer (1972, 1973) extended the Routley-Routley semantics to (full) relevant logics, they held that the real world must be among the worlds that has a ‘privileged status’ (Routley and Meyer (1973): 205). To identify a privileged world,<sup>7</sup> Routley and Meyer introduced non-normal worlds which do not have such a status. The introduction of non-normal worlds allowed them to invalidate irrelevant conditionals such as  $A \rightarrow (B \rightarrow B)$ .

At a non-normal world, the truth value of a conditional,  $A \rightarrow B$ , is not determined by the truth values of  $A$  and  $B$  at the same world. It is determined by those values at different worlds that are related in a ternary manner. What is important here is not the ternary relations but the fact that the antecedent and the consequent are evaluated at different worlds. That allows a non-normal world to make  $B \rightarrow B$  come out not true.

---

<sup>7</sup>There may be more than one privileged world. But completeness requires there to be only one such world.

But, as we saw above,  $B \rightarrow B$  is a logical truth in the system. Non-normal worlds are parts of the structure that validate  $B \rightarrow B$  as a logical truth. Yet it is at these worlds that  $B \rightarrow B$  fails to be true. So, a non-normal world is part of the structure that defines the logical truths of the system but it is a world where those very logical truths fail to hold.

It can now be shown that a non-normal world is a world where the laws of logic fail and is, thus, a logically impossible world. Consider a conditional  $A \rightarrow B$  where  $A$  and  $B$  do not contain any conditionals. The semantics for FDE is a systematic account of the truth-preserving relationship between the antecedent,  $A$ , and the consequent,  $B$ . It was developed to study the relationship between  $A$  and  $B$  in terms of  $A \models B$ . When it is extended to (full) relevant logics, the analysis of the relation between  $A$  and  $B$  is used to study the conditional  $A \rightarrow B$  in terms of  $\models A \rightarrow B$ . This means that if  $A \rightarrow B$  is a logical truth, the conditional expresses the validity of the truth-preserving relation between the antecedent and the consequent. In other words, in relevant logic, a logical truth in the form of conditional expresses a law of logic.

There is then a sense in which logical truths of conditional form express the laws of logic. And non-normal worlds introduced to the semantics of relevant logic are exactly the worlds where the laws of logic, understood as above, fail. They are logically impossible worlds. Thus, once we define logically impossible worlds in terms of non-normal worlds that Routley/Sylvan (with his collaborators) introduced to invalidate irrelevant conditionals such as  $A \rightarrow (B \rightarrow B)$ , we can gain a clear understanding of what makes a world logically impossible.

### 3 Non-Normal Worlds

There is nothing strange or mysterious about logically impossible worlds as I characterised them above. Such worlds are part of the semantic resources that Kripke (1963, 1965) introduced for the systems of modal logic developed by C.I. Lewis (especially for  $S2$  and  $S3$ ) and E.J. Lemmon (especially  $E2$  and  $E3$ ). This means that logically impossible worlds have been with us since the beginning of possible worlds semantics.<sup>8</sup>

Among the systems of modal logic are  $S2$  and  $S3$  as well as  $E2$  and  $E3$ , the so-called non-normal systems. The main characteristic of non-normal systems is the failure of necessitation: it is not the case that if  $\models A$  then  $\models \Box A$ . In order to develop semantics for these systems, Kripke introduced non-normal worlds where  $\Box A$  fails to hold for any  $A$ . Once we have these worlds, even if  $\models \Box(A \vee \neg A)$  (as  $A \vee \neg A$  is true at every world),  $\not\models \Box \Box(A \vee \neg A)$ .

---

<sup>8</sup>Material contained in this section draws on Tanaka (2013).

To make things more precise, let  $\langle W, N, R, \nu \rangle$  be a structure where  $W$  is a set of worlds,  $N$  is a set of *normal* worlds,  $R \subset W \times W$  and  $\nu$  is an evaluation of propositions at a world. For simplicity, we assume  $R$  to be universal. For any  $A$  and any  $w \in W - N$ ,  $\nu_w(\Box A) = 0$ .

Validity can be defined in two ways. First, it can be defined in terms of all *normal* worlds. Let's call such validity *weak* validity and represent it as  $\models_w$ . This gives rise to a Lewis non-normal system. Second, validity can be defined in terms of *all* worlds. Let's call it *strong* validity and represent it as  $\models_s$ . This gives rise to a Lemmon system of modal logic.<sup>9</sup>

Logical truths are special cases of validity. So  $A$  is a weak logical truth,  $\models_w A$ , if  $A$  is true at every normal world in every model. A strong logical truth,  $\models_s A$ , is defined in terms of all worlds. So  $\models_s A$  if  $A$  is true at every world in every model.

Let's define the *generality* of logic in terms of the claim that a logical truth expresses a truth no matter what the situation turns out to be. So to say that the necessity of  $A$  is general is to say that  $\Box A$  is a logical truth. Let  $\models A$  represent the generality of logical truth  $A$ . Then  $\models_s A$  expresses that the necessity of  $A$  is general, i.e.,  $\models_s A \Leftrightarrow \models \Box A$ .

If a law of logic is understood to be expressed by a logical truth, then  $\models \Box A$  means that  $\Box A$  is a law of logic. Thus, there is a sense in which laws of logic are expressed by logical truths of the form  $\Box A$ . If laws of logic are expressed by  $\Box A$ , then the failure of  $\Box A$  implies the failure of the laws of logic. So, the fact that  $\Box A$  fails to be true for any  $A$  invokes the idea that there are no laws of logic. Thus, a non-normal world introduced by Kripke can be characterised as a world where the laws of logic fail. Just like a non-normal world of relevant logic, it is a logically impossible world.

#### 4 Logically Impossible Worlds vs. Worlds Where the Laws of Logic are Different

Before we end our discussion about logically impossible worlds, let's consider an alternative characterisation that might be thought to define logically impossible worlds. Priest (2008) characterises a world where the laws of logic fail as synonymous with a world where the laws of logic are different. For instance, he writes:

non-normal worlds ... are worlds where the laws of logic are different.

Let us call these 'logically impossible worlds'. (p. 171)

A little bit later he talks about a logically impossible world as a world 'at which logical laws fail' (p. 172).

---

<sup>9</sup>Lemmon (1957). For a discussion, see Hughes and Cresswell (1996).

Priest elaborates on his characterisation of logically impossible worlds, worlds where the laws of logic are different, in analogous to physically impossible worlds. At a world where the laws of physics are different, objects might ‘accelerate past the speed of light’ (p. 172). According to Priest, at such a world, the laws of physics might permit physical objects to move differently from the physical laws of ‘our’ world. Nevertheless, at that world, there are laws of physics. It is just that they are different.

Now, Priest (1992) distinguishes the worlds where the laws of logic are different from the worlds where the laws of logic fail. However, he claims that the non-normal worlds which are instrumental for invalidating  $A \rightarrow (B \rightarrow B)$  are ‘worlds where the laws of logic are different, not worlds where the logically impossible happens’ (p. 296). Logically impossible worlds defined in this paper are, however, not the worlds where the laws of logic are different. What characterises them is the failure of the laws of logic. At a non-normal world, logical truths of the characteristic form (conditional in the case of relevant logics and necessary statement in the case of modal logics) fail to hold. However, it is not that different logical truths hold at a non-normal world. If  $A \rightarrow B$  expresses a law of logic, it is the failure of the conditional that defines a world being logically impossible. It does not entail, however, that the non-normal world is part of the structure that provides a different logical truth and the logical truth holds at that non-normal world. No laws of logic hold at a non-normal world. So a logically impossible world is a world where the laws of logic fail, but it is not a world where the laws of logic are different.

## References

- Berto, Francesco (2013) ‘Impossible Worlds’, *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), Edward N. Zalta (ed.), Stanford University.
- Cresswell, M. (1973) *Logics and Languages*, London: Methuen.
- Detlefsen, Michael (1990) ‘Brouwerian Intuitionism’, *Mind* 99: 501-534.
- Hughes, G.E., and M.J. Cresswell (1996) *A New Introduction to Modal Logic*, London: Routledge.
- Jago, Mark (2015) ‘Impossible Worlds’, *Noûs* 49: 713-728.
- Kripke, S.A. (1963) ‘Semantical Analysis of Modal Logic I. Normal Modal Propositional Calculi’, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 9: 67-96.
- (1965) ‘Semantical Analysis of Modal Logic II. Non-Normal Modal Propositional Calculi’, *The Theory of Models: Proceedings of the 1963 International Symposium at Berkeley*, J.W. Addison, L. Henkin and A. Tarski (eds.), Amsterdam: North-Holland Publishing Company: 206-20.
- Lemmon, E.J. (1957) ‘New Foundations for Lewis Modal Systems’, *The Journal of*

- Symbolic Logic* 22: 176-186.
- Lycan, W. (1994) *Modality and Meaning*, Dordrecht: Kluwer.
- Nolan, Daniel (1997) 'Impossible Worlds: A Modest Approach', *Notre Dame Journal of Formal Logic* 38: 535-572.
- (2013) 'Impossible Worlds', *Philosophy Compass* 8: 360-372.
- Priest, Graham (1992) 'What is a Non-Normal World?', *Logique et Analyse* 35: 219-302.
- (1997) 'Sylvan's Box', *Notre Dame Journal of Formal Logic* 38: 573-582.
- (2008) *An Introduction to Non-Classical Logic* (Second Edition), Cambridge: Cambridge University Press.
- and Sylvan, Richard (1992) 'Simplified Semantics for Basic Relevant Logics', *Journal of Philosophical Logic* 21: 217-232.
- Restall, Greg (1995) 'Four-Valued Semantics for Relevant Logics (and Some of Their Rivals)', *Journal of Philosophical Logic* 24: 139-160.
- Routley, R. (1989) 'Philosophical and Linguistic Inroads: Multiply Intensional Relevant Logics', J. Norman and R. Routley (eds.), *Directions in Relevant Logic*, Dordrecht: Kluwer: 269-304.
- Routley, R. and Meyer, R.K. (1972) 'The semantics of entailment – II', *Journal of Philosophical Logic* 1: 53-73.
- and Meyer, R.K. (1973) 'The semantics of entailment – I' *Truth, Syntax and Modality*, H. Leblanc (ed.), Amsterdam: North-Holland: 199-243.
- and Routley, V. (1972) 'The Semantics of First Degree Entailment', *Noûs* 6: 335-359.
- Tanaka, Koji (2013) 'Making Sense of Paraconsistent Logic: The Nature of Logic, Classical Logic and Paraconsistent Logic', *Paraconsistency: Logic and Applications*, K. Tanaka *et al.* (eds.), Dordrecht: Springer: 15-25.