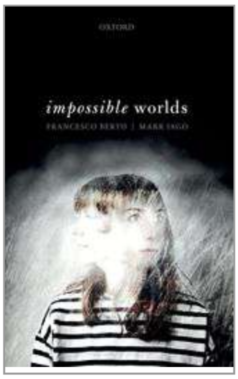

Impossible Worlds
Francesco Berto and Mark Jago

Print publication date: 2019
Print ISBN-13: 9780198812791
Published to Oxford Scholarship Online: August 2019
DOI: 10.1093/oso/9780198812791.001.0001

Modal Logics

Francesco Berto
Mark Jago

DOI:10.1093/oso/9780198812791.003.0004

Abstract and Keywords

This chapter introduces normal propositional modal logics, then non-normal systems which invalidate the Necessitation rule (N). It shows how to model these logics using non-normal or impossible worlds, thought of as 'logic violators'. This approach comes with non-uniform truth conditions: some operators are understood in one way at normal worlds, in another way at non-normal worlds. This may or may not be a problem. The specific case of non-adjunctive and non-prime worlds are then discussed, where conjunction and disjunction can behave in unusual ways.

Keywords: normal propositional modal logics, non-normal systems, Necessitation, non-normal worlds, non-adjunctive worlds, non-prime worlds

4.1 Normal Modal Logics

We begin with a rehearsal of standard or normal modal logic. To keep things simple, we limit ourselves to a propositional language $\mathscr{L}$, including a set of atoms *AT*: $p, q, r, p_1, p_2, \ldots$. We have negation ¬, conjunction ∧, disjunction ∨, the material conditional ⊃, and the box □ and diamond ◇ of necessity and possibility. We use $A, B, C, \ldots$ as metavariables for formulas of $\mathscr{L}$. The well-formed formulas are the atoms in *AT* and, if *A* and *B* are well-formed formulas, then so are:

$$\neg A | (A \land B) | (A \lor B) | (A \supset B) | \Box B | \Diamond B$$

Outermost brackets are normally omitted.

A *normal possible worlds frame* or *Kripke frame* $\mathscr{F}$ for $\mathscr{L}$ is a pair $\langle\langle W, R\rangle\rangle$, where W is a set of possible worlds and $R \subseteq W \times W$ is a binary accessibility relation between them. A frame becomes a *model* $\mathscr{M} = \langle\langle W, R, v\rangle\rangle$, when endowed with a valuation function v. This assigns to each atom either the value 1 (true) or the value 0 (false) at a world. So we write '$v_w(p) = 1$' to mean that p is true at w, and '$v_w(p) = 0$' to mean that it is false there.

The valuation function v is extended to the whole language via the following recursive clauses:

(S¬) $v_w(\neg A) = 1$ if $v_w(A) = 0$, and 0 otherwise.

(S∧) $v_w(A \wedge B) = 1$ if $v_w(A) = v_w(B) = 1$, and 0 otherwise.

(p.96)

(S∨) $v_w(A \vee B) = 1$ if $v_w(A) = 1$ or $v_w(B) = 1$, and 0 otherwise.

(S⊃) $v_w(A \supset B) = 1$ if $v_w(A) = 0$ or $v_w(B) = 1$, and 0 otherwise.

(S□) $v_w(\Box A) = 1$ if for all $w_1 \in W$ such that Rww_1, $v_{w_1}(A) = 1$, and 0 otherwise.

(S◇) $v_w(\Diamond A) = 1$ if for some $w_1 \in W$ such that Rww_1, $v_{w_1}(A) = 1$, and 0 otherwise.

Logical consequence or entailment '⊨', is defined as truth preservation at all worlds of all models (for any set of formulas Γ):

$\Gamma \vDash A$ iff for all models $\mathscr{M} = \langle\langle W, R, v\rangle\rangle$ and all $w \in W$: if $v_w(A) = 1$ for all $B \in \Gamma$, then $v_w(A) = 1$.

For single-premise entailment, we will write $A \vDash B$ instead of $\{A\} \vDash B$. Logical equivalence, $A \dashv\vDash B$, is two-way entailment between A and B. Logical validity or logical truth, $\vDash A$, defined as truth at all worlds of all models, is a special case of entailment by the empty set, $\varnothing \vDash A$.

The semantics makes $\Box A$ equivalent to $\neg\Diamond\neg A$ and $\Diamond A$ equivalent to $\neg\Box\neg A$, as desired. It also validates the *Distribution principle* or *K-principle*:

(K) $\Box(A \supset B) \supset (\Box A \supset \Box B)$

The logic induced by the semantics (the set of valid sentences) is called **K**, after Kripke. This is the weakest *normal modal logic*. In the context of modal logic, 'normal' means that the logic includes all the classical tautologies plus (K), and is closed under *modus ponens* and the *Necessitation* rule:

(N) If $\vdash A$, then $\vdash \Box A$

(Be careful in how you read this rule. It doesn't say that A implies $\Box A$: that would trivialize modality by committing us to treating all truths as necessary truths. Rather, it says that, if A is a theorem of the logic/a logical truth, then so is $\Box A$.)

(p.97) **K** is the base normal modal logic, in that its semantics puts no conditions on the accessibility relation R. If we impose some conditions on R, we obtain stronger normal modal logics. The normal modal logics obtained in this way contain all the **K**-theorems, plus some extra ones too. (In semantic terms, we get more entailments by putting further conditions on the accessibility relation R.) Table 4.1 shows the most well-known cases:

Table 4.1: axiom-frame correspondence

Axiom name	Axiom scheme	Frame condition
D	$\Box A \supset \Diamond A$	R is serial: $\forall x \exists y Rxy$
T	$\Box A \supset A$	R is reflexive: $\forall x Rxx$
B	$A \supset \Box\Diamond A$	R is symmetrical $\forall x \forall y (Rxy \rightarrow Ryx)$
4	$\Box A \supset \Box\Box A$	R is transitive $\forall x \forall y \forall z (Rxy \wedge Ryz \rightarrow Rxz)$
5	$\Diamond A \supset \Box\Diamond A$	R is euclidean $\forall x \forall y \forall z (Rxy \wedge Rxz \rightarrow Ryz)$

The logic **KTB** adds the T and B axioms, for example, and so corresponds to to the reflexive and symmetrical frame. Corresponding to the serial and euclidean frame is the logic **KD5**. And so on.

Whether we accept these additional axioms depends on how we understand the involved modalities. The D axiom says that what is necessary is possible, and this seems plausible on most readings of the notions of possibility and necessity. For D to hold, we need every world in W to access some world (that's Seriality). For if a world w accesses no world, given (S$\Box$), all formulas of the form $\Box A$ are true at w. And given (S$\Diamond$), all formulas of the form $\Diamond A$ are false at w, for there is no accessible world where A is true. Thus, D fails. (The label 'D' comes from 'deontic', inspired by the reading of necessity as 'it ought to be the case that' and of possibility as 'it is permissible that'.)

If the relevant necessity is factive, then the T axiom must hold: if A is necessary (at a given world w), then it should be true (at that **(p.98)** world). This won't hold in general without Reflexivity (every world is possible relative to itself). For without Reflexivity, it could be that A holds at all worlds accessible from w but not at w itself.

If the relevant necessity is unrestricted, then the 4 and 5 axioms look very plausible. If A is unrestrictedly necessary (or possible) in the relevant sense, *this* fact should not be contingent on anything and so it should be necessary as well. This may not be so for factive but restricted or relative necessities. Against 4, for instance, it might be physically necessary (determined by the laws of physics) that bodies do not accelerate through the speed of light; but its necessity may not be determined by the laws of physics. The 4 and 5 axioms are characteristic of two important normal modal logics, **S4** (= **KT4**) and **S5** (= **KT5**), due to C.I. Lewis (Lewis and Langford 1932), a founding father of modern modal logic.

If we read '$\Box$' as an epistemic operator expressing knowledge, it is doubtful that either 4 or 5 holds. In epistemic logic, 4 is called the Axiom of (Positive) Introspection, or *KK-principle*. It says that, if one knows that A, then one knows that one knows that A. One has perfect introspective access to what one knows. It seems, however, that this has counterexamples. Think of yourself panicking the night before the exam, but doing fine with your essay the day after. You may truthfully say: 'Yesterday I didn't know I had learned so much by studying, but today it turned out that I did'. You knew the answers, but didn't know that you knew them all.

The 5 axiom is even more suspect in an epistemic setting. It is equivalent to $\neg\Box A \supset \Box\neg\Box A$, which, in an epistemic setting, says: if one doesn't know that A, then one knows that one doesn't know that A. That doesn't seem at all plausible. For one thing, we often think we know things we don't in fact know. In those cases, we don't believe, and so don't know, that we don't know them.

As we shall see in Chapter 5, it is indeed doubtful that *any* normal modal logic can provide an adequate formal treatment of epistemic notions like knowledge and belief. This is due to the logical omniscience phenomena, which were introduced in Chapter 1, and to which we shall return in §5.1.

(p.99) 4.2 Non-Normal Modal Logics

This section expands on Berto and Jago 2018. Normal Kripke frames are celebrated for having provided suitable interpretations of different systems of modal logic, including **S4** and **S5**. Before Kripke's work, we merely had lists of axioms or, at most, algebraic semantics many found rather uninformative. Kripke also introduced *non-normal worlds* (Kripke 1965), in order to provide world-based semantics for modal logics weaker than the basic normal modal system **K**. These are *non-normal* modal logics, including C.I. Lewis's systems **S2** and **S3**.

Non-normal modal systems do not include the Necessitation rule:

(N) If $\vdash A$, then $\vdash\Box A$

As we said in §4.1, (N) holds in the weakest normal modal logic **K** and all its normal extensions. A semantic counterpart of (N) would tell us that if A is a logical truth, then so is $\Box A$. This principle cannot be avoided when $\Box$ is understood in line with (S$\Box$). For if A is a logical truth, then it is by definition true at all worlds of all models. So given any world w, A is true at all worlds accessible from w, so also $\Box A$ is true at w. Since this applies to any world of any model, $\Box A$ will thereby be a logical truth, too.

Non-normal worlds enter the stage in order to make (N) fail. Take the same language $\mathscr{L}$ of §4.1 and give it the following semantics. A *non-normal worlds frame* $\mathscr{F}$ for $\mathscr{L}$ is a triple $\langle\langle W, N, R\rangle\rangle$, with W the set of worlds and $N \subseteq W$ the subset of normal worlds, so that the items in $W - N$ are the non-normal worlds. R is as before. A frame becomes a non-normal model $\mathscr{M} = \langle\langle W, N, R, v\rangle\rangle$ when endowed with a valuation function v assigning truth values to formulas at worlds.

The truth conditions for the extensional logical vocabulary are defined as in §4.1. But we now take the clauses (S$\Box$) and (S$\Diamond$) to apply to normal worlds only. If $w \in W - N$, the clauses are:

(NS$\Box$) $v_w(\Box A) = 0$

(NS$\Diamond$) $v_w(\Diamond A) = 1$

(p.100) At non-normal worlds, formulas of the form $A, with $ a modal operator, are not evaluated depending on the truth value of A at other (accessible) worlds, but get assigned their truth value directly. Specifically, all $\Box$-formulas are false and all $\Diamond$-formulas are true. In a sense, non-normal worlds of this kind are worlds where nothing is necessary and anything is possible. These worlds are deviant only in this respect: their behavior, as far as the extensional connectives are concerned, is quite regular. Notice also that, as is easy to check, (NS$\Box$) and (NS$\Diamond$) still deliver the equivalence of $\Box A$ with $\neg\Diamond\neg A$ and of $\Diamond A$ with $\neg\Box\neg A$.

Logical consequence or entailment is defined as truth preservation at all *normal* worlds in all models:

$\Gamma \vDash B$ iff for all models $\mathscr{M} = \langle\langle W, N, R, v\rangle\rangle$ and all $w \in N$: if $v_w(A) = 1$ for all $A \in \Gamma$, then $v_w(B) = 1$

Logical validity is truth at all normal worlds in all models.

Restricting logical consequence and validity to normal worlds in this way is a common, though not universal, move in semantics that include non-normal or impossible worlds. The insight behind this comes from the second characterization of impossible worlds as 'logic violators' from §1.4: worlds where logic is different, or where the laws of logic fail. If this is the interpretation of the items in $W - N$, then we should not refer to, or quantify over, such worlds when we characterize logical consequence and validity. For these, we want to look only at possible or normal worlds: worlds where logic is *not* different.

This setting gives us a basic non-normal modal logic, which Priest (2008) calls **N**. If one adds the condition that R be reflexive, one gets C.I. Lewis's modal system **S2**. If one takes R to be reflexive and transitive, one gets **S3** (Kripke 1965).

This kind of semantics makes (N) fail. Take any classical propositional tautology, say, of the form $A \vee \neg A$. This holds at all worlds of all interpretations. Therefore, $\Box(A \vee \neg A)$ holds at all normal worlds of all interpretations, so $\vDash \Box(A \vee \neg A)$. But by (NS$\Box$), $\Box(A \vee \neg A)$ does **(p.101)** not hold in any non-normal world. So $\Box\Box(A \vee \neg A)$ is false at normal worlds that have access, via R, to any non-normal world. As there must be some such world in some model, we have $\nvDash \Box\Box(A \vee \neg A)$.

Another welcome feature of this semantics is that it does not make the 'irrelevant' conditional $A \dashv (B \dashv B)$ valid. This is one of the 'paradoxes' of the strict conditional (§1.3). It fails in non-normal models once one reads $A \dashv B$ as the necessitation of the horseshoe, $\Box(A \supset B)$. Then the paradox is $\Box(A \supset \Box(B \supset B))$. This is valid in **K**, but not in **N** (and extensions): take a world $w \in N$ accessing a non-normal w_1 where A holds but (since the world is non-normal) $\Box(B \supset B)$ fails.

Non-normal worlds semantics of this kind does not provide a systematic framework for dealing with all cases of irrelevance (nor was it intended to do so). Nevertheless, this way of handling one paradox of the strict conditional hints at a general strategy: make irrelevant conditionals fail by taking into account non-normal or impossible worlds, understood as worlds where logical truths such as $B \dashv B$ may fail. We'll see more of this approach in Chapter 6.

The semantics for non-normal modal logics such as **S2** and **S3** is based on a valuation function which assigns the same truth value to all $\Box$-formulas (false) and all $\Diamond$-formulas (true) at non-normal worlds. We also mention the modal system **S0.5**, due to Lemmon (1957). This is a non-normal system whose semantics, initially provided by Cresswell (1966), includes non-normal worlds at which formulas that begin with a modal operator are assigned arbitrary truth values. The valuation function v treats modal formulas as atomic. (Interpretations for **S2** or **S3** are special cases of interpretations for **S0.5**: those cases in which the valuation function uniformly treats $\Box$-formulas as false and $\Diamond$-formulas as true at non-normal worlds.) This setting makes the inter-definability of $\Box$ and $\Diamond$ via negation fail.

(p.102) 4.3 Non-Uniform Truth Conditions

A key idea in impossible worlds semantics of various kinds is that certain complex formulas are assigned arbitrary values at non-normal worlds. They are, in effect, treated as atomic sentences. At worlds that behave this way, the syntax of a formula can be partly or wholly disregarded. As we shall see in Chapter 5 on epistemic logics, this insight can be fruitful, and is at work in semantic frameworks including non-normal or impossible worlds.

This approach requires the truth conditions of some operators not to be spelled out in a uniform way across worlds. In particular, at non-normal worlds complex formulas can have truth values assigned in a non-recursive way. But this approach raises worries. Fine (2019) claims that it is a 'theoretical virtue in itself' for a semantics to be uniform:

> we would like the compositional clauses for the logical connectives to be 'uniform' or non-disjunctive. … without uniformity, it is not even clear that we will have clauses for the logical connectives themselves as opposed to some gerry-mandered product of the theoretician's mind.
>
> Fine (2019, 1)

A similar worry is pressed in Williamson (2017).

It is not clear, however, why disjunctiveness would be a problem. The fact that a concept has a disjunctive characterization per se does not make the concept itself gerrymandered or gruesome. The notion *Australian citizen*, for instance, is obviously perfectly fine even though it works just as follows: *x* is Australian citizen iff either *x* was born in Australia or *x* has been naturalized. (The example is due to Priest (2005, 237).)

A more serious worry is lack of compositionality. If truth-at-a-world-conditions constitute meaning, and we want meanings to be compositional for them to be graspable by finite minds, then the truth-at-a-world-conditions of whole formulas should be given in terms of those of their subformulas. If there are infinitely many non-normal worlds in our setting, however, we may have no way finitely to do this.

(p.103) We think this is a serious philosophical worry. We'll offer a philosophical response to it in §8.5, by showing how to get a compositional account of content involving impossible (non-normal) worlds. Here, our focus is on the logical applications of impossible worlds. We want our impossible worlds models to give us a notion of logical consequence that's useful for certain applications, for example. But that isn't to demand that they are all compositional. A model can be useful, even if it isn't capable of underpinning a theory of meaning for a language like English. So for pragmatic logical purposes, we're happy to dismiss the worry.

4.4 Non-Adjunctive and Non-Prime Worlds

This section draws on Berto and Jago 2018. Rescher and Brandom (1980) introduce impossible worlds of a different kind to the ones we just saw. (Their book's subtitle is *A Study in Non-Standard Possible-Worlds Semantics and Ontology*, but, as we shall now see, the worlds making their semantics non-standard deserve the label of impossible worlds.)

In their formal semantics, standard possible worlds are taken as maximal consistent aggregates of states of affairs. Their non-standard worlds are obtained combinatorially, by means of two recursive operations having standard worlds as their base. These are *schematization* ('∩') and *superposition* ('∪'). Given w_1 and w_2, a schematic world $w_1 \cap w_2$ is one at which all and only the states of affairs obtain which obtain both at w_1 and at w_2. A superposed (or inconsistent) world $w_1 \cup w_2$ is one at which all and only the states of affairs obtain which obtain either at w_1 or at w_2. With respect to the definitions listed in §1.4, Rescher and Brandom's inconsistent-superposed worlds are impossible worlds of the fourth kind, that is, contradiction-realizers: they can make both A and $\neg A$ true. (Just superpose a possible world, w_1, at which you are 1.70m tall, and another one, w_2, at which you are not.)

(p.104) Unlike Kripke's non-normal worlds, which require different truth conditions only for the modal operators, Rescher and Brandom's non-standard worlds behave peculiarly with respect to extensional operators. The standard clause for conjunction (S∧) from §4.1 has to go for superposed worlds: A and B can each be true without $A \wedge B$'s being true. These worlds still have a certain amount of logical structure. They behave in quite a standard fashion with respect to essentially single-premise inferences (Priest et al. 1989, 161). But they are anarchic with respect to essentially multiple-premise inferences.

Dually, schematic worlds can be *non-prime*: it can happen that $A \vee B$ is true at a world without either A or B being true at that world. One application that can motivate non-prime worlds is the handling of under-determined information: one may have the information that Strasbourg is either in France or in Germany, without having information as to which is the case.

Now 'dualizing' back, one may use inconsistent-superposed worlds to model inconsistent databases. These may consist in sets of data or information, supplied by different sources which are inconsistent with each other, such as incompatible evidence presented by different witnesses in a trial. The non-adjunctive features of superposed worlds are useful here. Intuitively, one is allowed to draw the logical consequences of the data or information fed in by a single source, but one does not conjoin data from distinct sources which may be inconsistent with each other. The database is 'compartimentalized': occasional inconsistencies are placed in separate sectors, and not conjunctively asserted. This is an example of a *non-adjunctive* system. Hyde (1997), Lewis (1982), and Varzi (1997) each discuss different uses of this kind of approach. We will come back to uses of impossible worlds in information theory in more detail in Chapter 9.

Chapter Summary

After recapping standard normal modal logics and their frame correspondeces (§4.1), we introduced non-normal modal logics, **(p.105)** which invalidate the Necessitation rule (N). We showed how to model these logics using non-normal or impossible worlds, thought of as ‘logic violators’ (§4.2). This approach gives comes with non-uniform truth conditions: some operators are understood in one way at normal worlds, in another way at non-normal worlds (§4.3). This may or may not be a problem; we’ll come back to it in §8.5. We then discussed the specific case of non-adjunctive and non-prime worlds (§4.4), where conjunction and disjunction can behave in unusual ways. **(p.106)**