# The Metaphysics of Routley Star*

Edward N. Zalta
Philosophy Department
Stanford University

**Abstract**

This paper investigates two forms of the Routley star operation, one in Routley & Routley 1972 and the other in Leitgeb 2019. We use the background of object theory to define the first form of the operation and show that, under a minimal additional assumption, the second form can be defined as well. Then we show that the principles governing both forms become derivable and need not be stipulated. Since no mathematics is assumed by our background formalism, the existence of the Routley star image $s^*$ of a situation $s$ is therefore guaranteed not by set theory but by a theory of abstract objects. The work in the paper integrates Routley star into a more general theory of (partial) situations that has previously been used to develop the theory of possible worlds and impossible worlds.

## 1 Introduction

The Routley 'star' operation was introduced in Routley & Routley 1972. Their study of the semantics of entailment assumed the existence of situations ('set-ups') that are neither consistent nor maximal (*ibid.*, 335–339).[1] When the Routleys set up the star operator on situations, they used '$H$' to range over set-ups (i.e, "a class of propositions or wff") and used '$A$' to range over propositions or wffs (*ibid.*, 337). Then they considered the following condition (*ibid.*, 338) on the star (*) operation, which they labeled as (iv):

(iv) $\sim A$ is in $H$ iff $A$ is not in $H^*$

They subsequently stipulated that a set-up is $\sim$-*normal* if it satisfies (iv) for every $A$ and $H = H^{**}$ (*ibid.*, 338).

That was then. Although the Routley star has been studied and applied in a number of subsequent works, it was recently used in Leitgeb 2019 (321ff) to build a semantics for a system of hyperintensional logic ('HYPE'). Leitgeb first builds a propositional language $\mathcal{L}$ that includes propositional letters, with some standard logical connectives, but with a non-standard conditional. Leitgeb then constructs HYPE-models for $\mathcal{L}$ in terms of structures whose elements include a non-empty set of states $S$ and a valuation function $V$ from $S$ to the power set of the set of literals of the language $\mathcal{L}$, so that each state $s$ in $S$ is associated with a set of literals $V(s)$. I'll describe HYPE models in fuller detail below, but for the purposes of this introduction, it is important to note that the various elements of HYPE models are simultaneously constrained by the requirements of a Routley star operation having the following properties, among others (Leitgeb 2019, 322):

- $V(s^*) = \{\overline{v} \mid v \notin V(s)\}$
- $s^{**} = s$

Leitgeb then discusses the properties of the star operation and uses HYPE models to define various truth conditions for hyperintensional operators.

These two bookend cases, Routley & Routley 1972 and Leitgeb 2019, demonstrate how the Routley star operation has been deployed to help us understand various non-classical, but more fine-grained, semantic phenomena. But a look at the body of literature inclusive between these papers, a metaphysician would find that relatively little attention has been paid to the question: What kind of metaphysics is represented by a

*The research in this paper was first sketched for presentation in Hannes Leitgeb's seminar *Logic and Metaphysics*, which was held at the Munich Center for Mathematical Philosophy in May 2022. I subsequently developed the results into a section of *Principia Logico-Metaphysica* (Zalta, m.s.). I'm indebted to Hannes Leitgeb, Uri Nodelman, Daniel Kirchner, Daniel West, Graham Priest, and an anonymous referee for their comments about this material, all of which helped me to refine and improve some of the results.

[1] Some logicians use the term 'non-normal worlds' to describe situations that are neither maximal (complete) nor consistent. The Routleys, however, used the term 'world' for consistent and maximal situations (1972, 339). In what follows, we reserve the term 'world' for maximal situations, some of which are possible worlds and some of which are impossible worlds.

semantics making use of Routley star, and how are we to understand the Routley star operation given that metaphysics?

Questions about the meaning of the Routley star operation were raised early on, in Copeland 1979 and van Benthem 1979. Restall 1999 (54) raises this question when he wrote:

> The operator $^*$ was introduced to relevant logic by Routley and Routley [23]. If $x \neq x^*$, then certainly we can get both $A \wedge {\sim}A \rightarrow B$ and $A \rightarrow B \vee {\sim}B$ to fail, but there is a price. The price is the obligation to explain the meaning of the operator $^*$.

But even though we may now be more comfortable with Routley star and recognize how interesting and efficacious it is (given the work that has been done), there is still an open question about what, exactly, is the proper metaphysical framework for defining and studying the Routley star operation.

In our two case studies, and for most studies in between, one typically finds the Routley star introduced into semantic models constructed with the help of set theory, domains of primitive entities (set-ups, situations, states, possible or impossible worlds), and functions defined on those domains, etc. Most authors don't spend time considering the metaphysics of the entities used in their semantic models, and quite rightly, given their goals. For their purposes, it is sufficient to adopt another attitude expressed in Restall 1999 (57):

> It would be interesting to chart the connections between states as we have sketched them and other entities like ... objects, states of affairs, propositions, and many other things besides. However, this is neither the time nor the place for that kind of metaphysics. Suffice it to say that a coherent comprehensive view of states ought to tell us how these things fit together. For now, we will use *states* as the points in our frames for relevant logics.

For example, Leitgeb writes (2019, 323, footnote 9):

> I want to leave open in this paper whether states are interpreted (i) in a metaphysically robust manner, or (ii) in a looser informational manner. In the first case, states would be "chunks of reality" that are "located in the world", while in the second case they might be some kind of abstract entities corresponding to "pieces of thought".



By contrast, Mares 2004 (4.4–4.11) attempts to develop an intuitive understanding of the assumptions concerning properties, states of affairs, situations, propositions, etc., that are used in the semantic models. But (a) the focus of Mares 2004 is to interpret the ternary relation $R$ used in Routley-Meyer semantics for relevant logic (Routley & Meyer 1972, 1973), and (b) Mares assumes that some background theory of situations is available, such as Barwise and Perry 1983, for he takes a number of principles about situations as given.

By contrast, in what follows, we plan to develop the metaphysics of Routley star without any mathematics, set theory, primitive domains of situations, states, or worlds (possible or impossible), or functions on domains. We won't identify propositions as sets of possible worlds, as functions from possible worlds to truth values, as sets of situations, or as classes of wffs. Nor will we assume any axioms governing primitive set-ups, situations, possible worlds, or impossible worlds. Instead, we shall *define* the Routley star operator metaphysically in *object theory* (= OT), where situations are defined and their first principles derived. And we employ a theory of propositions (= 0-ary relations) that is part of a larger, hyperintensional theory of $n$-ary relations – one on which necessarily equivalent relations and propositions aren't identified. Basic OT allows us to define a unique Routley star image $s^*$ for each situation $s$, as in Routley & Routley 1972. Moreover, with a minimal, additional assumption (or axiom, if you prefer), $s^*$ can be defined as in Leitgeb 2019. Our goals, then are to show that, in such a setting, (a) the metaphysical entities needed to formulate and understand the Routley star image can be defined and proved to exist, (b) the principles governing Routley star, as formulated in both Routley & Routley 1972 and Leitgeb 2019, can be *derived* rather than stipulated, and (c) the connections between the two definitions of Routley star can be precisely articulated.

The present effort may be distinguished from other recent discussions of Routley star by its methodology and the focus of the investigation. Relatively recent papers such as Restall 2000, Berto 2015, and Berto & Restall 2019 (all of which acknowledge a debt to Dunn 1993) are about the semantic analysis of various forms of negation and, as such, the action takes place in the semantics. In each of these papers, a frame semantics involving a primitive relation of *compatibility* (or *incompatibility*) on *points* or primitive *worlds* is introduced and used to interpret an uninterpreted language with a negation symbol.[2] By contrast, it is *not* a goal of the present paper to study or define non-classical negation semantically. The reader will find no semantics in what follows. Rather, situations, possible worlds, impossible worlds, and the Routley star are all defined. The definitions are cast in a logic and metaphysics that is systematized proof-theoretically. Moreover, notions very much like the notions of incompatibility and compatibility utilized in the papers by Restall and Berto will also be *defined*, and the key principles that govern them will be derived rather than stipulated.[3]

Another distinguishing feature of the present work is its stated goal of reconciling Routley star as developed Routley & Routley 1972 and Leitgeb 2019.[4] So by investigating the metaphysics of Routley star in the manner below, the present effort may help us better understand the domain of application for Routley star and thereby better prepare us for understanding the uses to which it has been put in the semantics of non-classical negation, both in the two papers that serve as the focus of our study and in other papers on the semantics of negation.

[2]In Restall 2000, semantic frames and a primitive relation of compatibility on points are introduced on the first page. In Berto 2015, frames are introduced (766ff), and negation is analyzed as a modality (767) that is interpreted by a distinguished accessibility relation on worlds, $R_N$, understood as a compatibility relation (768ff). In Berto & Restall 2019, the semantic analysis occurs in Section 3, where frames and the primitive compatibility relation on worlds are introduced (1127).

[3]See Section 4.3 for the definition of *incompatibility* and see footnote 18 for (a) a way to define the *compatibility* relation that Restall and Berto take as primitive, and (b) a derivation of the principle they stipulate to characterize that relation (Restall 2000, 853, Definition 1.1; Berto 2015, 768, 'Backward'; and Berto & Restall 2019, 1129, 'Backwards').

A second principle, the Heredity Principle (Restall 2000, Definition 1.2; Berto 2015, 767; and Berto & Restall 2019, 1128), was previously derived in OT as the *Persistence* principle (Zalta 1993, 413, Theorem 8); this settled a choice point in Barwise 1989 (265) in favor of Alternative 6.1.

Finally, see footnote 15 below for a discussion of how the reflexivity, anti-symmetry, and transitivity principles governing the relation $\sqsubseteq$ on the points of compatibility frames, stipulated in Restall 2000 (853, Definition 1.1), was previously derived in OT (Zalta 1993, 413, Theorem 7) in terms of the condition $s \trianglelefteq s'$ on object-theoretic situations.

[4]Papers published subsequent to Leitgeb 2019 have had other goals. See Odintsov & Wansing 2020 for a comparison of the hyperintensional propositional logic in HYPE with a number of other logics, and Punčochář & Sedlár 2022 for a discussion of the Routley star operation in information-based semantics rather than truth-conditional semantics.



## 1.1 The Background Theory

The background theory needed to achieve these goals has been motivated and published elsewhere (see below) and we shall draw on those published results. In what follows, the reader should be familiar with the fact that OT is expressed in a second-order, quantified modal language (without identity) that includes two kinds of atomic formulas: standard *exemplification* formulas of the form $F^n x_1 \ldots x_n$ and *encoding* formulas of the form $xF$. This language is extended with complex individual terms, namely (rigid) definite descriptions of the form $\imath x\varphi$, and with complex $n$-ary relation terms of the form $[\lambda x_1 \ldots x_n\, \varphi]$ ($n \geq 0$). A primitive unary predicate $E!$ (*being concrete*) is used to distinguish *ordinary* objects ($O!x \equiv_{df} \diamond E!x$) and *abstract* objects ($A!x \equiv_{df} \neg\diamond E!x$). Identity for objects is defined: $x = y$ holds if and only if either $x$ and $y$ are both ordinary objects that necessarily exemplify the same properties or both abstract objects that necessarily encode the same properties.

The underlying logic of OT includes: (a) classical propositional logic, (b) classical predicate logic for the constants and variables but negative free logic for the complex terms (i.e., descriptions and $\lambda$-expressions may fail to denote), and (c) full S5 modal logic, including the Barcan (1946) and converse Barcan formulas for both the first and second-order variables (i.e., there are fixed domains of objects and relations). A $\lambda$-expression of the form $[\lambda x_1 \ldots x_n\, \varphi]$ is guaranteed to have a denotation if none of the variables bound by the $\lambda$ occur as an argument term in an encoding formula in $\varphi$. $\lambda$-Conversion (aka $\beta$-Conversion) holds for such $\lambda$-expressions, as does $\alpha$-Conversion.[5] Identity is also *defined* for properties, relations, and propositions and the definitions do *not* imply that necessarily equivalent properties, relations, or propositions are identical.[6] Finally, there are three axioms in OT that govern the logic of encoding. One asserts that ordinary objects necessarily fail to encode

[5]For *elementary* $\lambda$-expressions of the form $[\lambda x_1 \ldots x_n\, Fx_1 \ldots x_n]$, $\eta$-Conversion holds. It asserts $[\lambda x_1 \ldots x_n\, Fx_1 \ldots x_n] = F$, where $F$ is an $n$-ary relation variable.

[6]The key definition is for the identity of properties, i.e., unary relations (Zalta 1993, 407)):

$$F = G \equiv_{df} \Box\forall x(xF \equiv xG)$$

Then, 0-ary relations (i.e., propositions) $p$ and $q$ are defined to be identical just in case the propositional properties $[\lambda x\, p]$ and $[\lambda x\, q]$ are identical (Zalta 1993, 409). And $n$-ary relations $F$ and $G$ ($n \geq 2$) are identical just in case each way of 'plugging' $n-1$ objects into both $F$ and $G$ yields identical properties. We omit the last two formal definitions here, since they won't be needed.

properties ($O!x \rightarrow \Box\neg\exists F xF$). A second asserts that the modal logic of encoding is rigid ($\Diamond xF \rightarrow \Box xF$). And the main axiom is a comprehension schema for abstract objects, which asserts that for any condition $\varphi$ with no free $x$s, there is an abstract object that encodes all and only the properties such that $\varphi$:

$$\exists x(A!x \,\&\, \forall F(xF \equiv \varphi)), \text{ provided } x \text{ isn't free in } \varphi \tag{1}$$

Further details of the system will be brought to bear as the occasion arises.

In Zalta 1993 and 1997, OT was deployed to develop the theory of situations, possible worlds, and impossible worlds. The theory begins with the definitions:

- a *situation* is any abstract object that encodes only properties of the form *being such that* $p$ (i.e., properties of the form $[\lambda x\, p]$, where $x$ is vacuously bound by the $\lambda$, and $p$ is a variable ranging over propositions):

$$\mathit{Situation}(x) \equiv_{df} A!x \,\&\, \forall F(xF \rightarrow \exists p(F = [\lambda x\, p])) \tag{2}$$

- $p$ *is true in situation* $s$ (‘$s \models p$’), or $s$ *makes* $p$ *true*, is defined as $s$ encodes the propositional property *being such that* $p$:

$$s \models p \equiv_{df} s[\lambda x\, p] \tag{3}$$

In OT, ‘$\models$’ always takes the *smallest* scope; so $s \models p \rightarrow p$ is to be parsed $(s \models p) \rightarrow p$; otherwise, we write $s \models (p \rightarrow p)$. Also, we sometimes read $s \models p$ as $s$ encodes $p$, thereby extending the notion of encoding.

In Zalta 1993 (410–414), it was shown that the basic principles of situation theory are derivable from the definition of *situation* given above. Indeed, 15 of the 19 principles outlined in Barwise 1989 were derived. World theory is then an extension of situation theory and it developed via the following definition:

- a *possible world* is any situation $s$ that might be such that all and only true propositions are true in $s$:

$$\mathit{PossibleWorld}(s) \equiv_{df} \Diamond\forall p(s \models p \equiv p) \tag{4}$$

  Given our convention, the subformula $s \models p \equiv p$ is to be parsed as $(s \models p) \equiv p$.



- an *impossible world* is any maximal situation (i.e., for every proposition $p$, either $s$ makes $p$ true or $s$ makes the $\neg p$ true) for which it is not possible that every proposition true in $s$ is true:

$$\mathit{Maximal}(s) \equiv_{df} \forall p(s \models p \vee s \models \neg p) \tag{5}$$

$$\mathit{ImpossibleWorld}(s) \equiv_{df} \mathit{Maximal}(s) \,\&\, \neg\Diamond\forall p(s \models p \rightarrow p) \tag{6}$$

The basic principles of possible world theory are derivable from the definition of *possible world* given above (Zalta 1993, 414–419). These include formal versions of the following principles:

- every possible world is maximal, consistent, and modally closed;
- there is a unique actual world;
- possibly $p$ iff there is a possible world in which $p$ is true; and
- necessarily $p$ iff $p$ is true in every possible world.

And the basic principles of impossible world theory can be derived from the definition of *impossible world* given above (Zalta 1997, 646–649). These include formal versions of:

- there are impossible worlds;
- if it is not possible that $p$, then there exists a non-trivial impossible world in which $p$ is true;[7]
- there exist impossible worlds where the principle *ex contradictione quodlibet* (ECQ) fails; and
- there exist impossible worlds where disjunctive syllogism fails.

The above principles were all shown to be theorems. Familiarity with the foregoing results will be presupposed in what follows, since we now plan to extend and build upon them.

[7] Cf. Nolan (1997, 542), who similarly suggests that impossible worlds are governed by the comprehension principle: for every proposition that cannot be true, there is an impossible world where that proposition is true.

## 1.2 The Recent Developments We'll Need

Among the recent developments of OT we'll need for the analysis of Routley star are the following definition and theorem schema:

$$\overline{p} =_{df} \neg p \tag{7}$$

$$\vdash \exists s \forall p(s \models p \equiv \varphi), \text{ provided } s \text{ isn't free in } \varphi \tag{8}$$

Definition (7) lets us denote the negation of a proposition more simply as $\overline{p}$. As a theorem schema, (8) is in fact a comprehension schema for situations and is derivable from axiom (1). A derivation of (8) is given in the Appendix. It is also provable that situations $s$ and $s'$ are identical just in case they make the same propositions true (Zalta 1993, 412):

$$\vdash s = s' \equiv \forall p(s \models p \equiv s' \models p) \tag{9}$$

Consequently, it follows immediately from (8) that there is a unique situation that makes true all and only the propositions satisfying $\varphi$:

$$\vdash \exists! s \forall p(s \models p \equiv \varphi), \text{ provided } s \text{ isn't free in } \varphi \tag{10}$$

It is a consequence of (10) that every definite description having the form $\imath s \forall p(s \models p \equiv \varphi)$ is always well-defined (i.e., provably has a denotation), provided $s$ isn't free in $\varphi$. These are, therefore, *canonical* descriptions for situations.

## 1.3 Some Other Non-classical Situations

Earlier we described how OT implies the existence of impossible worlds in which certain classical laws of logic fails to hold. But it is important to remember, as we work through the results below, that we don't have to consider impossible worlds to find situations in which the laws of classical logic fail. Classical laws may fail in situations that aren't impossible worlds. Consider the law *ex contradictione quodlibet* (ECQ) and let $q_1$ be any proposition. Then the following is an instance of (10), which asserts the existence of a unique situation that makes exactly one proposition true, namely, the conjunction $q_1 \,\&\, \neg q_1$:

$$\exists! s \forall p(s \models p \equiv p = (q_1 \,\&\, \neg q_1))$$

Call this situation $s_1$, so that we know $\forall p(s_1 \models p \equiv p = (q_1 \,\&\, \neg q_1))$. Clearly, the conjunction $q_1 \,\&\, \neg q_1$ is true in $s_1$, i.e., $s_1 \models (q_1 \,\&\, \neg q_1)$. Now consider



any proposition that is distinct from the conjunction $q_1 \,\&\, \neg q_1$, say, $r_1$. It then follows that $\neg s \models r_1$. So we've established that for any propositions $q$ and $r$ such that $(q \,\&\, \neg q) \neq r$, there is a situation in which ECQ fails:

$$\vdash \forall q \forall r\big((q \,\&\, \neg q) \neq r \to \exists s(s \models (q \,\&\, \neg q) \,\&\, \neg s \models r)\big)$$

Indeed, we can use the above to define a condition that isolates precisely those situations that fail ECQ:

$$s \textit{ is an ECQ-falsifier} \equiv_{df} \exists q \exists r(s \models (q \,\&\, \neg q) \,\&\, \neg s \models r)$$

Similarly, we can define a group of situations in which disjunctive syllogism (DS) fails. Let $q_1$ and $r_1$ be any propositions such that the propositions $q_1 \vee r_1$, $\neg q_1$, and $r_1$ are all pairwise distinct. Then consider following instance of (10), which asserts the existence of a unique situation that encodes exactly two propositions, namely, $q_1 \vee r_1$ and $\neg q_1$:

$$\exists! s \forall p\big(s \models p \equiv p = (q_1 \vee r_1) \vee p = \neg q_1\big)$$

Call this $s_2$, so that we know $\forall p\big(s_2 \models p \equiv p = (q_1 \vee r_1) \vee p = \neg q_1\big)$. Then it is easy to establish all of the following: $s_2 \models (q_1 \vee r_1)$, $s_2 \models \neg q_1$, and $\neg s_2 \models r_1$. So DS fails with respect to $s_2$. And, in general, we have established that for any pairwise distinct propositions $q \vee r$, $\neg q$, and $r$, there is a situation in which DS fails:

$$\vdash \forall q \forall r\big((q \vee r) \neq \neg q \,\&\, \neg q \neq r \,\&\, (q \vee r) \neq r) \to \exists s(s \models (q \vee r) \,\&\, s \models \neg q \,\&\, \neg s \models r)\big)$$

Again, we can use the above to define a group of situations in which DS fails:

$$s \textit{ is a DS-falsifier} \equiv_{df} \exists q \exists r(s \models (q \vee r) \,\&\, s \models \neg q \,\&\, \neg s \models r)$$

These examples are of interest because they show that OT already has the capacity to develop counterexamples to classical logical laws without Routley star, once those classical laws are interpreted within the domain of situations. We don't need to formulate a separate language and define truth for the formulas of that language with respect to the domain of situations. The condition $s \models \varphi$ (i.e., $\varphi$ is *true in* $s$) is defined for all situations $s$ and formulas $\varphi$. That's because *every* formula $\varphi$ denotes a proposition,[8] and so each $\varphi$ (with no free $x$s) can be instantiated for $p$

[8]To see this, recall that in OT, propositions are 0-ary relations. So let $\Pi$ be an arbitrary 0-ary relation term. In the latest developments of object theory (Zalta m.s.), we define:

in definition (3) to obtain an instance of the form $s \models \varphi \equiv_{df} s[\lambda x\, \varphi]$.[9] So our notion of *true in* situation $s$ applies to arbitrary formulas and we can directly evaluate the truth of formulas relative to any distinguished (i.e., definable) group of situations.

Moreover, one can define, for example, the *conjunction-normal* situations as those situations that make $p\,\&\,q$ true whenever they both make $p$ true and make $q$ true. Formally:

$$s \textit{ is conjunction-normal} \equiv_{df} \forall p \forall q \big(s \models (p \,\&\, q) \equiv (s \models p \,\&\, s \models q)\big)$$

And $s$ is *double-negation normal* just in case $s$ makes $\overline{\overline{p}}$ true if and only if it makes $p$ true. And so on. One may therefore precisely define, for some particular application, which group of situations is to be studied.

## 1.4 Canonical Descriptions and Modality

At the end of Section 1.2, we identified *canonical* descriptions of the form $\imath s \forall p(s \models p \equiv \varphi)$. Though canonical descriptions are always logically proper, one must take care when deploying them in a modal context, given that, in OT, the formal definite description $\imath x \varphi$ rigidly denotes the unique object, if there is one, that satifies $\varphi$ at the distinguished actual world. It is worth digressing a moment to understand the issues that arise and why the present paper will be able ignore them. We conclude the digression and this section by formulating a theorem schema involving descriptions that will play an important role in the paper.

Note that in a modal logic with rigid definite descriptions, one can produce logical theorems that are not necessary. For example, the conditional $y = \imath x Gx \to Gy$ will be false at a world, say $w_1$, when $y$ (is assigned an object that) fails to be $G$ at $w_1$ but is the unique $G$ at the actual world $w_0$ (in such a case, the the antecedent is true at $w_1$ but the consequent false at $w_1$). More generally, where $\varphi^y_x$ is the result of substituting $y$ for all the free occurrences of $x$ in $\varphi$, the claim $y = \imath x \varphi \to \varphi^y_x$ is not a necessary truth, though it is logically true (i.e., true at the distinguished actual world of every model, for every assignment to $x$) given the semantics of rigid definite descriptions.

In a fuller presentation of OT, we could axiomatize rigid definite descriptions by introducing an actuality operator $\mathcal{A}$ and asserting, as an axiom:

$$y = \imath x \varphi \equiv \forall x(\mathcal{A}\varphi \equiv x = y) \tag{11}$$

This is a form of the Hintikka principle (1959); it is a necessary truth and it immediately implies the following as a necessary truth, in which $\mathcal{A}\varphi^y_x = (\mathcal{A}\varphi)^y_x = \mathcal{A}(\varphi^y_x)$:

$$\vdash y = \imath x \varphi \to \mathcal{A}\varphi^y_x, \text{ provided } y \text{ is substitutable for } x \text{ in } \varphi \tag{12}$$

If we then adjust the original example, it should be easy to see that $y = \imath x Gx \to \mathcal{A}Gy$ is a necessary truth. But though (11), (12), and their instances are necessary truths, the axiomatization of the actuality operator includes an axiom, namely $\mathcal{A}\varphi \to \varphi$, that is a logical truth which isn't necessary (Zalta 1988).[10] So the Rule of Necessitation has to be slightly adjusted; one may not apply the rule to necessitate a theorem whose proof depends on the axiom $\mathcal{A}\varphi \to \varphi$.

In what follows, though, we won't need to worry about illicit applications of the Rule of Necessitation since all of the definite descriptions we'll deploy involve a special class of formulas for which we can derive the conditional $y = \imath x \varphi \to \varphi^y_x$ without appealing to the contingent axiom for actuality. The formulas in question are *modally collapsed*, i.e., any formula $\varphi$ for which it is provable that $\Box(\varphi \to \Box\varphi)$. When a formula having this form is provable, one can prove $\mathcal{A}\varphi \equiv \varphi$ without appealing to the contingent axiom $\mathcal{A}\varphi \to \varphi$.[11] If $\varphi$ is modally collapsed, then $y = \imath x \varphi \to \varphi^y_x$ is a necessary truth:

---

$\Pi{\downarrow}$ (read: $\Pi$ exists) just in case $[\lambda \nu\, \Pi]{\downarrow}$, where $\nu$ is some variable not free in $\Pi$. But the definiens, which asserts that the propositional property $[\lambda \nu\, \Pi]$ exists, is axiomatic, since it meets the condition that the bound variable $\nu$ doesn't occur as an argument in an encoding formula anywhere in $\Pi$. So, it is provable that $\Pi{\downarrow}$. But, in OT, formulas are 0-ary relation terms and since $\Pi$ was arbitrary, it is a theorem of OT that $\varphi{\downarrow}$, for any formula $\varphi$. So every formula denotes a proposition.

[9] The restriction that $x$ not be free in $\varphi$ is no real restriction. If $\varphi$ has a free variable $x$, then choose a variable that is not free in $\varphi$. Without loss of generality, suppose it is $y$. Then as an instance of an alphabetic variant of definition (3), we have $s \models \varphi \equiv_{df} s[\lambda y\, \varphi]$. So the definition holds for any formula $\varphi$.

[10] To see why the formula schema $\mathcal{A}\varphi \to \varphi$ can't be necessitated, note that the conditional is true at the actual world: if $\varphi$ is true at the actual world, then the conditional is true at the actual world (by truth of the consequent), and if $\varphi$ is false at the actual world, then the conditional is true at the actual world (by failure of the antecedent). However, the conditional is false at any world $w_1$ whenever $\varphi$ is true at the actual world but false at $w_1$.

[11] Assume $\Box(\varphi \to \Box\varphi)$. Then by the K$\Diamond$ principle, i.e., $\Box(\psi \to \chi) \to (\Diamond\psi \to \Diamond\chi)$, it follows that $\Diamond\varphi \to \Diamond\Box\varphi$. But in S5, $\Diamond\Box\varphi \to \Box\varphi$. So by hypothetical syllogism, we've established:

($\theta$) $\Diamond\varphi \to \Box\varphi$

Now to see that $\mathcal{A}\varphi \equiv \varphi$, we prove both directions. ($\to$) Assume $\mathcal{A}\varphi$. Then $\Diamond\varphi$. So by ($\theta$),

$$\vdash y = \imath x\varphi \to \varphi^y_x, \tag{13}$$
provided $\varphi$ is modally collapsed and $y$ is substitutable for $x$ in $\varphi$

(See the Appendix for the proof.) In this paper, we shall appeal only to definite descriptions in which the matrix is modally collapsed, and so we won't need to worry about mistakenly applying the Rule of Necessitation to theorems derived from a logical truth that is not necessary.

In particular, we have, as a special case of (13), that when $\varphi$ is modally collapsed, then if a situation $s$ is identical to the situation that makes true all and only the propositions satisfying $\varphi$, then $s$ makes true all and only the propositions satisfying $\varphi$, i.e.,

$$\vdash s = \imath s'\forall p(s' \models p \equiv \varphi) \to \forall p(s \models p \equiv \varphi), \tag{14}$$
provided $s'$ isn't free in $\varphi$ and $\varphi$ is modally collapsed

The keys to the proof in the Appendix are the facts that $s' \models p$ is, by definition (3), an instance of the formula $xF$ and that the modal logic of encoding is $xF \to \Box xF$. So by the Rule of Necessitation, $\Box(xF \to \Box xF)$ and, as an instance, $\Box(s' \models p \to \Box s' \models p)$. This fact, and the fact that $\varphi$ is modally collapsed, lets us validly infer that the formula $\forall p(s' \models p \equiv \varphi)$ is modally collapsed. So the description $\imath s'\forall p(s' \models p \equiv \varphi)$ will be governed by (13).

(14) plays a crucial role in what follows. All of descriptions of the form $\imath s'\forall p(s' \models p \equiv \varphi)$ used in the present work will be constructed in terms of formulas $\varphi$ that are modally collapsed; it is provable that their truth necessarily implies their own necessity. This should forestall any concerns about the fact that we shall be working within a modal context in which definite descriptions are interpreted rigidly.

## 2 Definitions and Theorems

For any situation $s$, we define *the Routley star image of* $s$, written $s^*$, as the situation $s'$ that makes true all and only those propositions whose negations aren't true in $s$:

$$s^* =_{df} \imath s'\forall p(s' \models p \equiv \neg s \models \overline{p}) \tag{15}$$

$\Box\varphi$. Hence $\varphi$, by the T schema. ($\leftarrow$) Assume $\varphi$. Then $\Diamond\varphi$. But again by ($\theta$), it follows that $\Box\varphi$. Hence $\mathcal{A}\varphi$.



Clearly, the definiens has a denotation: it is a canonical description for which $s'$ doesn't occur free in $\neg s \models \overline{p}$. So $s^*$ is well-defined. Since it can be shown that $\neg s \models \overline{p}$ is a modally collapsed formula, it follows from (15) by (14) that $p$ is true in $s^*$ iff $\overline{p}$ fails to be true in $s$:

$$\vdash \forall p(s^* \models p \equiv \neg s \models \overline{p}) \tag{16}$$

This holds for any situation $s$. (The first part of the proof in the Appendix establishes that $\neg s \models \overline{p}$ is a modally collapsed formula.)

We now establish a number of facts that show (15) and theorem (16) properly capture the definition of $s^*$ in Routley & Routley 1972. Since formulas of the form $\varphi \equiv \neg\psi$ are necessarily equivalent to formulas of the form $\neg\varphi \equiv \psi$, (16) implies that, for any proposition $p$, $\overline{p}$ is true in $s$ if and only if $p$ fails to be true in $s^*$:

$$\vdash \forall p(s \models \overline{p} \equiv \neg s^* \models p) \tag{17}$$

Again, this holds for any situation $s$. (17) is an analogue of the Routleys' principle (iv), as formulated in the opening paragraph of Section 1 above.

To set up the next confirmation that (15) is correct, let us say that $s$ *has a glut with respect to* $p$, written $GlutOn(s,p)$, if and only if both $p$ and $\overline{p}$ are true in $s$; and that $s$ *has a gap with respect to* $p$, written $GapOn(s,p)$, if and only if neither $p$ nor $\overline{p}$ is true in $s$:

$$GlutOn(s,p) \equiv_{df} s \models p \;\&\; s \models \overline{p} \tag{18}$$

$$GapOn(s,p) \equiv_{df} \neg s \models p \;\&\; \neg s \models \overline{p} \tag{19}$$

Then it follows that the condition $s = s^{**}$ implies that if $s$ has a glut with respect to $p$, then $s^*$ has a gap with respect to $p$:

$$\vdash s = s^{**} \to (GlutOn(s,p) \to GapOn(s^*,p)) \tag{20}$$

And $s = s^{**}$ also implies that if $s = s^{**}$, then if $s$ has a gap with respect to $p$, then $s^*$ has a glut with respect to $p$:

$$\vdash s = s^{**} \to (GapOn(s,p) \to GlutOn(s^*,p)) \tag{21}$$

Moreover, it can be shown, without the assumption that $s = s^{**}$, that if $s$ neither has a glut nor a gap w.r.t. $p$, then $s^*$ makes $p$ true if and only if $s$ makes $p$ true:

$$\vdash (\neg GlutOn(s,p) \;\&\; \neg GapOn(s,p)) \to (s^* \models p \equiv s \models p) \tag{22}$$

It then follows that if, for every proposition $p$, $s$ neither has a glut nor a gap w.r.t. $p$, then $s^* = s$ (since they make the same propositions true); and for every proposition $p$, $s$ neither has a glut nor a gap w.r.t. $p$ if and only if for every proposition $p$, $s$ makes $p$ true if and only if $s$ fails to make $\overline{p}$ true:

$$\vdash \forall p(\neg GlutOn(s,p) \,\&\, \neg GapOn(s,p)) \rightarrow s^* = s \tag{23}$$

$$\vdash \forall p(\neg GlutOn(s,p) \,\&\, \neg GapOn(s,p)) \equiv \forall p(s \models p \equiv \neg s \models \overline{p}) \tag{24}$$

Intuitively, (24) tells us that if $s$ is free of gluts and gaps, then it is coherent with respect to negation.

We conclude this section by deriving three interesting facts, the first two of which require us to define *the null situation* ($\boldsymbol{s}_{\emptyset}$), in which no propositions are true, and *the trivial situation* ($\boldsymbol{s_V}$), in which every proposition is true:

$$\boldsymbol{s}_{\emptyset} =_{df} \imath s' \forall p(s' \models p \equiv p \neq p) \tag{25}$$

$$\boldsymbol{s_V} =_{df} \imath s' \forall p(s' \models p \equiv p = p) \tag{26}$$

The facts are that: if $s^{**} = s$ holds universally, then the Routley star image of the null situation is the trivial situation; if $s^{**} = s$ holds universally, then the Routley star image of the trivial situation is the null situation; and $s^{**}$ is identical to $s$ if and only if, for every proposition $p$, $p$ is true in $s$ iff $\overline{\overline{p}}$ is true in $s$:

$$\vdash \forall s(s^{**} = s) \rightarrow \boldsymbol{s}_{\emptyset}{}^* = \boldsymbol{s_V} \tag{27}$$

$$\vdash \forall s(s^{**} = s) \rightarrow \boldsymbol{s_V}{}^* = \boldsymbol{s}_{\emptyset} \tag{28}$$

$$\vdash s^{**} = s \equiv \forall p(s \models p \equiv s \models \overline{\overline{p}}) \tag{29}$$

(29) becomes interesting when we consider the passage in Routley & Routley 1972 (338) in which they discuss their principle (iv), which we reproduced above in the opening paragraph of Section 1:

> Requirement (iv) on its own does not suffice for the normality of the negation, since it does not assume such characteristic negation features as double negation features. For these features it is, however, unnecessary to adopt the over-restrictive condition $H = H^*$, which would take us back to (ii); it suffices to require that $H = H^{**}$.



The Routleys don't say here exactly which double negation features they are referring to. But (29) tells us that the condition $s^{**} = s$ is equivalent to a specific double negation feature. As we've seen, the Routleys go on to suggest that a 'set-up', i.e., a situation $s$, is classical ('normal') w.r.t. double negation when $s^{**} = s$. Even if the fact expressed by (29) has been made explicit somewhere else in the literature, it has now been derived from general principles that don't assume any mathematics, and the derivation occurs in a purely logical and metaphysical system in which propositions have been axiomatized, and situations and their Routley star images have been defined.

## 3 An Alternative Definition

In Section 5 below, we investigate an alternative definition of the Routley star image. Instead of defining $s^*$ as the situation that makes true all and only the propositions whose negations aren't true in $s$, the alternative defines $s^*$ as the situation that makes true all and only the negations of propositions that aren't true in $s$:

$$s^* =_{df} \imath s' \forall p(s' \models p \equiv \exists q(\neg s \models q \,\&\, p = \overline{q})) \tag{$\vartheta$}$$

Since the condition $\exists q(\neg s \models q \,\&\, p = \overline{q})$ is modally collapsed,[12] ($\vartheta$) immediately implies, by (14):

$$\forall p(s^* \models p \equiv \exists q(\neg s \models q \,\&\, p = \overline{q})) \tag{$\xi$}$$

($\vartheta$) and its consequence ($\xi$) are of interest because the key condition, $\exists q(\neg s \models q \,\&\, p = \overline{q})$, is *not* equivalent to the condition $\neg s \models \overline{p}$ used in (15); in other words, ($\vartheta$) and (15) don't always define the same $s^*$ for any given situation $s$.[13]

[12] We can show this by first noting that both conjuncts of this quantified conditional are modally collapsed. Since $s \models q$ is modally collapsed (see the discussion immediately following (14), so is $\neg s \models q$. Moreover, as noted in footnote 6, $p = q$ holds just in case the property identity $[\lambda x\, p] = [\lambda x\, q]$ holds, where the identity of properties $F = G$ is defined as $\Box\forall x(xF \equiv xG)$. Given the S4 axiom then, it is easy to show $F = G \rightarrow \Box F = G$. So by the Rule of Necessitation $\Box(F = G \rightarrow \Box F = G)$. Instantiating $F$ and $G$ to $[\lambda x\, p]$ and $[\lambda x\, q]$ and applying the definition of identity for propositions, we have the instance $\Box(p = q \rightarrow \Box p = q)$, which holds for any propositions $p$ and $q$. Hence $\Box(p = \overline{q} \rightarrow \Box p = \overline{q})$.

So it remains to show that the quantified conjunction is modally collapsed. But if $\varphi$ and $\psi$ are modally collapsed, it follows that $\varphi \,\&\, \psi$ is modally collapsed, i.e., if $\Box(\varphi \rightarrow \Box\varphi)$ and $\Box(\psi \rightarrow \Box\psi)$, then $\Box((\varphi \,\&\, \psi) \rightarrow \Box(\varphi \,\&\, \psi))$. From these facts it doesn't take much more work to show $\Box\big(\exists q(\neg s \models q \,\&\, p = \overline{q}) \rightarrow \Box\exists q(\neg s \models q \,\&\, p = \overline{q})\big)$.

[13] In what follows, it is important to distinguish the following two conditions:

To see why, consider a simple situation, say $s_1$, in which a single proposition, say $p_1$, is true. Let's ignore all other propositions and consider what propositions are true in $s_1^*$ according to (16) vs. what propositions are true $s_1^*$ according to consequence ($\xi$). According to (16), the following propositions are true in $s_1^*$:

- $p_1$ (since $\neg s_1 \models \overline{p_1}$),
- $\overline{p_1}$ (since $\neg s_1 \models \overline{\overline{p_1}}$),
- $\overline{\overline{p_1}}$ (since $\neg s_1 \models \overline{\overline{\overline{p_1}}}$),
- and so on.

But according to ($\xi$), neither $p_1$ nor $\overline{p_1}$ are true in $s_1^*$ (neither $p_1$ nor $\overline{p_1}$ is the negation of a proposition that $s_1$ *fails* to encode). Instead, the following propositions are true in $s_1$ according to ($\xi$):

- $\overline{\overline{p_1}}$ (since $\neg s_1 \models \overline{p_1}$ and $\overline{\overline{p_1}}$ is the negation of $\overline{p_1}$),
- $\overline{\overline{\overline{p_1}}}$ (since $\neg s_1 \models \overline{\overline{p_1}}$ and $\overline{\overline{\overline{p_1}}}$ is the negation of $\overline{\overline{p_1}}$),
- and so on.

So the two alternative ways of defining $s_1^*$, namely (15) and ($\vartheta$), yield different situations even in this very simple case. Clearly, then, one can't simply replace the definiens of (15) with the definiens:

$$\imath s' \forall p(s' \models p \equiv \exists q(\neg s \models q \,\&\, p = \overline{q}))$$

This won't preserve the results we've established thus far.

Though there may be multiple ways one could bring (15) and ($\vartheta$) into alignment and force them into defining the same Routley star situation for a given $s$, the simplest way is to limit some of the hyperintensionality in propositions. In particular, we can show that $\exists q(\neg s \models q \,\&\, p = \overline{q})$ and $\neg s \models \overline{p}$ become equivalent whenever propositions are identical to their double negations, i.e., whenever:

(1) $\exists q(\neg s \models q \,\&\, p = \overline{q})$

(2) $\exists q(\neg s \models q \,\&\, q = \overline{p})$

Condition (2) is equivalent to $\neg s \models \overline{p}$, by the following argument:

> ($\rightarrow$) Assume $\exists q(\neg s \models q \,\&\, q = \overline{p})$ and suppose $r$ is such a propositions, so that we know both $\neg s \models r$ and $r = \overline{p}$. Then $\neg s \models \overline{p}$. ($\leftarrow$) Assume $\neg s \models \overline{p}$. Then $\neg s \models \overline{p} \,\&\, \overline{p} = \overline{p}$, by the reflexivity of identity and &I. Hence, $\exists q(\neg s \models q \,\&\, q = \overline{p})$.

But we're now going to focus on condition (1), to see why it isn't equivalent to $\neg s \models \overline{p}$.



$$\forall p(\overline{\overline{p}} = p) \qquad (\zeta)$$

Consider how ($\zeta$) plays a role in the proof of both directions of the biconditional asserting the following equivalence:

$$\exists q(\neg s \models q \,\&\, p = \overline{q}) \equiv \neg s \models \overline{p} \qquad (\omega)$$

> *Proof*: ($\rightarrow$) Assume $\exists q(\neg s \models q \,\&\, p = \overline{q})$ and let $r$ be such a proposition, so that we know both $\neg s \models r$ and $p = \overline{r}$. The latter implies that $\overline{p} = \overline{\overline{r}}$, for if propositions are identical, so are their negations. But by ($\zeta$), $\overline{\overline{r}} = r$. Hence, $\overline{p} = r$ and so $\neg s \models \overline{p}$. ($\leftarrow$) Assume $\neg s \models \overline{p}$. Then by ($\zeta$), $\neg s \models \overline{p} \,\&\, p = \overline{\overline{p}}$. By existentially generalizing on $\overline{p}$ we have: $\exists q(\neg s \models q \,\&\, p = \overline{q})$. ⋈

Note that OT does *not* imply ($\zeta$) since the identity conditions of relations and propositions are hyperintensional; one may consistently claim that propositions and their double negations are distinct despite being necessarily equivalent. That's because in OT, propositions $p$ and $q$ are identical just in case the corresponding propositional properties $[\lambda x\, p]$ and $[\lambda x\, q]$ are identical, where property identity is, in turn, defined in terms of being necessarily encoded by the same objects, not in terms of being necessarily exemplified by the same objects (see footnote 6). Consequently, necessarily equivalent properties and propositions are not identified; properties and propositions are more fine-grained. But one can easily and consistently add ($\zeta$) as an axiom to OT or, in the alternative, derive consequences from the assumption that ($\zeta$) and thereby derive conditional theorems in which ($\zeta$) is the antecedent, via the Deduction Theorem.

But we don't even have to go as far as adding ($\zeta$) an axiom. In Section 4.1, we'll (a) define a group of propositions that are identical with their double negations, (b) assert only that there are at least some such propositions (without asserting that every proposition is identical to its double negation), and then (c) focus our attention on situations that are built out of such propositions. Then, in Section 5, we'll use ($\vartheta$) to define Routley star relative to that group of situations.

# 4 HYPE

Leitgeb (2019, 321ff) builds a semantics for a system of hyperintensional propositional logic ('HYPE'). He first builds a propositional language $\mathcal{L}$

by starting with atomic propositional letters $p_1, p_2, \ldots$, and logical symbols $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\top$ (where $\rightarrow$ does not express the material conditional). He writes $\overline{p_i}$ for $\neg p_i$, and uses $\overline{\overline{p_i}}$ as an abbreviation for $p_i$. The proposition letters and their negations constitute the *literals*. Leitgeb then constructs HYPE-models for $\mathcal{L}$ in terms of structures $\langle S, V, \circ, \perp \rangle$, where the elements of the models are simultaneously constrained by the requirements of a Routley star operation $^*$. He describes the elements of the models as follows (Leitgeb 2019, 321–22):

- $S$ is a non-empty set of states.
- $V$ is a function (the valuation function) from $S$ to the power set of the set of literals of the language $\mathcal{L}$, so that each state $s$ in $S$ is associated with a set of literals $V(s)$.
- $\circ$ is a partial fusion function on states that is idempotent and, when defined, commutative and partially associative.
- $\perp$ is a relation of incompatibility that relates states $s$ and $s'$ when some proposition $p$ is true at one and its negation $\overline{p}$ is true at the other. [Note: The *relation* $\perp$ among HYPE states is not to be confused with the symbol $\perp$ that Leitgeb uses as a metalinguistic abbreviation of the proposition $\neg\top$ (2019, 321). In what follows, the context should make it clear whether $\perp$ denotes the relation or the proposition.]

The Routley star operation that constrains these models will be discussed and defined later, in Section 5.

Consequently, in the remainder of this paper, we use OT to reconstruct the above elements of HYPE models and we'll see that the reconstruction comports with both of the suggestions for understanding HYPE states quoted above in Leitgeb 2019 (323, footnote 9). In Section 4.1 we develop basic definitions and show how to interpret the HYPE $V$ function; in Section 4.2 we show how to interpret the HYPE fusion operation $\circ$; and in Section 4.3, we show how to interpret the HYPE incompatibility relation $\perp$. Finally, in Section 5, we define the HYPE version of Routley star and prove that it has the expected features.



### 4.1 HYPE Propositions and HYPE States

First, we work our way towards a definition of a *Hype*-state by defining *Hype*-propositions. We say that (30) a *Hype*-proposition is any proposition $p$ that is identical to its double negation:

$$Hype(p) \equiv_{df} \overline{\overline{p}} = p \tag{30}$$

Clearly, then it follows that (31) if $p$ is a *Hype*-proposition, then so is its negation $\overline{p}$:

$$\vdash Hype(p) \rightarrow Hype(\overline{p}) \tag{31}$$

Though OT guarantees the existence of propositions (by 0-ary relation comprehension) and provides identity conditions for them (footnote 12), it doesn't guarantee the existence of *Hype*-propositions. The identity conditions for propositions in OT leave one free to assert the existence of *Hype*-propositions and the existence of propositions that are more fine-grained, e.g., by asserting $\exists p(\overline{\overline{p}} \neq p)$. Though $\Box(\overline{\overline{p}} \equiv p)$ is a theorem, it doesn't follow that $\overline{\overline{p}} = p$.

However, it is a trivial matter to extend OT by asserting the existence of at least some *Hype*-propositions. This hypothesis shouldn't be controversial to logicians who have worked in systems where propositions are represented as functions from possible worlds to truth values; such logicians have implicitly accepted that *all* propositions are identical with their double negations in those systems, since such propositions have the same truth value at every possible world. Nor does such a hypothesis significantly increase the resources OT needs for its analysis; it only ensures the existence of a subclass of propositions from among those that already exist. More importantly, by asserting there are *Hype*-propositions, we are asserting a metaphysical counterpart of Leitgeb's stipulation that the symbol '$\overline{\overline{p}}$' is to be an abbreviation of '$p$' (i.e., when he defines the HYPE language; 2019, 321). Every proposition letter in HYPE's language is thereby identified with its double negation, but here we require only that some propositions are so identifiable.

Moreover, we need not even assert this hypothesis as an axiom; it is sufficient to take it on board as an assumption. Of course, by asserting the hypothesis as an axiom, the results below would all become theorems. But to accomplish the goals of this paper, we need only show *what follows* in OT from the assumption that there are *Hype*-propositions. This demonstrates that if we extend OT with a principle ($\exists p Hype(p)$) used to frame the target logic (HYPE), the result is a metaphysical system for defining HYPE's Routley star and deriving the principles that govern it. By collapsing at least some propositions and their negations, and deriving the basic principles of HYPE relative to those propositions, we establish that one can analyze the metaphysics of HYPE in terms of a domain of propositions that are not themselves hyperintensional entities; the hyperintensionality will arise via other means.

So, in what follows, we shall work under the assumption that there are *Hype*-propositions:

$$\exists p Hype(p) \tag{32}$$

Then we may define $x$ is a *HypeState* just in case $x$ is a situation such that every proposition true in $x$ is a *Hype*-proposition:

$$HypeState(x) \equiv_{df} Situation(x) \,\&\, \forall p(x \models p \rightarrow Hype(p)) \tag{33}$$

So we're identifying *HypeStates* not as primitive entities but as situations. Thus when Leitgeb speaks of the members of $V(s)$ as the facts or states of affairs obtaining at $s$ (2019, 322), we may interpret this in terms of our defined notion, *p is true in s*, as follows:

- $p \in V(s) \equiv_{df} s \models p$

Now it is easy to prove the existence of *HypeStates*; (32) guarantees there are *Hype*-propositions and (8) guarantees that for any condition on *Hype*-propositions, there are situations that make true all and only such propositions. Clearly, any such situation is a *HypeState*.

Indeed, we now derive, from (8), comprehension conditions for *HypeStates* with the help of some new variables. Note that the conditions *Hype*($p$) and *HypeState*($x$), defined respectively in (30) and (33), are modally collapsed conditions. So may use introduce rigid restricted variables to range over them. For clarity, we use special new variables in a distinguished, sans-serif font:

- $\mathsf{p}, \mathsf{q}, \ldots$ are restricted variables ranging over *Hype*-propositions.
- $\mathsf{s}, \mathsf{s}', \ldots$ be are restricted variables ranging over *HypeStates*.

Using these variables we may formulate Simplified Comprehension for *HypeStates* as follows:

$$\vdash \exists \mathsf{s} \forall \mathsf{p}(\mathsf{s} \models \mathsf{p} \equiv \varphi), \text{ provided } \mathsf{s} \text{ isn't free in } \varphi \tag{34}$$



Clearly, in the usual way, it is provable that there is a unique such *HypeState* for each such instance:

$$\vdash \exists! \mathsf{s} \forall \mathsf{p}(\mathsf{s} \models \mathsf{p} \equiv \varphi), \text{ provided } \mathsf{s} \text{ isn't free in } \varphi \tag{35}$$

And this yields *canonical* descriptions for *HypeStates* that are always well-defined:

$$\vdash \imath \mathsf{s} \forall \mathsf{p}(\mathsf{s} \models \mathsf{p} \equiv \varphi)\downarrow, \text{ provided } \mathsf{s} \text{ isn't free in } \varphi \tag{36}$$

Clearly, then, there are *HypeStates* that are counterexamples to classical laws, just like the ones described in Section 1.3. Those show us how to construct *HypeStates* that falsify Explosion (ECQ) and Disjunctive Syllogism, i.e., two of the laws mentioned in the last bullet point of Observation 14 in Leitgeb 2019 (333). It should be clear how to construct *HypeStates* that falsify some of the others mentioned there as well, such as Excluded Middle, Law of Non-Contradiction, and (General) Contraposition.[14] However, one of the two examples mentioned in Observation 14, of a HYPE-invalid inference involving the proposition $\bot$ (namely, $\varphi \rightarrow \bot \models \neg\varphi$), will be discussed in the next section.

## 4.2 The HYPE Fusion Operation

If we put aside, for the moment, the fact that the fusion function $\circ$ in HYPE is a *partial* binary operation on *HypeStates*, then we can represent the HYPE $\circ$ operation as the following (total) *summation* operation $\oplus$ on situations generally:

$$s \oplus s' \;=_{df}\; \imath s'' \forall p\big(s'' \models p \equiv (s \models p \vee s' \models p)\big) \tag{37}$$

In other words, $s \oplus s'$ is the situation that makes a proposition $p$ true just in case either $s$ makes $p$ true or $s'$ makes $p$ true. Since $s \models p \vee s' \models p$ is modally collapsed, it follows that a proposition $p$ is true in $s \oplus s'$ just in case either $p$ is true in $s$ or $p$ is true in $s'$:

$$\vdash \forall p\big(s \oplus s' \models p \equiv (s \models p \vee s' \models p)\big) \tag{38}$$

[14] For example, let $\mathsf{p}_1$ be an arbitrary *Hype*-proposition and suppose $\mathsf{s}_1$ is the *HypeState* that makes just $\mathsf{p}_1$ true and $\neg\mathsf{p}_1$ true (and no other *Hype*-propositions true). Then it does not follow in OT that $\mathsf{s}_1 \models (\mathsf{p}_1 \vee \neg\mathsf{p}_1)$. So Excluded Middle doesn't hold in $\mathsf{s}_1$: it is not provable that $\mathsf{s} \models (\varphi \vee \neg\varphi)$ for arbitrary $\varphi$. And any *HypeState* $\mathsf{s}$ that isn't maximal will be such that there are *Hype*-propositions $\mathsf{q}$ such that neither $\mathsf{s} \models \mathsf{q}$ nor $\mathsf{s} \models \neg\mathsf{q}$. We leave it to the reader to construct *HypeStates* in which the Law of Non-Contradiction and (General) Contraposition are false.

To see that $\oplus$ captures additional features about the partial nature of situations generally, let us say that *s is a part of* $s'$ just in case every proposition true in $s$ is true in $s'$:[15]

$$s \trianglelefteq s' \equiv_{df} \forall p(s \models p \rightarrow s' \models p) \tag{39}$$

It follows relatively straightforwardly that $s$ is a part of $s'$ if and only if the sum of $s$ and $s'$ just is $s'$:

$$\vdash s \trianglelefteq s' \equiv s \oplus s' = s' \tag{40}$$

A further consequence of these definitions and theorems is that $\oplus$ is idempotent, commutative, and associative with respect to situations *generally*. Since *HypeStates* are situations, it follows that:

$$\vdash \oplus \text{ is idempotent, commutative, and associative on } \textit{HypeStates}. \tag{41}$$

Formally:

$$\vdash \mathsf{s} \oplus \mathsf{s}$$
$$\vdash \mathsf{s} \oplus \mathsf{s}' = \mathsf{s}' \oplus \mathsf{s}$$
$$\vdash \mathsf{s} \oplus (\mathsf{s}' \oplus \mathsf{s}'') = (\mathsf{s} \oplus \mathsf{s}') \oplus \mathsf{s}''$$

So by ignoring the partiality of $\circ$, we may interpret $s \circ s'$ in Leitgeb 2019 as $\mathsf{s} \oplus \mathsf{s}'$.

But the $\circ$ operation is in fact partial while $\mathsf{s} \oplus \mathsf{s}'$ is defined for *any HypeStates* $\mathsf{s}$ and $\mathsf{s}'$. So one might be concerned that our results don't properly capture the metaphysics of the fusion operation in HYPE. I don't think that concern is justified, however. For one thing, we could in fact model the partiality of $\circ$ by introducing a partial ternary relation $R^3$ (*not* the ternary relation $R$ of Routley-Meyer 1972, 1973) that may or may not

[15] The definition that follows was derived as a theorem in Zalta 1993 (412), as a consequence of the more general definition $x \trianglelefteq y \equiv_{df} \forall F(xF \rightarrow yF)$ and the fact that situations encode only propositional properties. But for the present investigation, we may simply take the following as a definition. It follows from our definition that $\trianglelefteq$ is reflexive, anti-symmetric, and transitive on situations generally (Zalta 1993, 413, Theorem 7), and on HypeStates in particular (exercise). Compare these theorems about $\trianglelefteq$ with Barwise 1989b (185) and 1989a (259), where they are taken as axioms of situation theory. Similarly, Restall 2000 (853), stipulates that there is an analogous relation $\sqsubseteq$ that is reflexive, anti-symmetric and transitive on the points of *compatibility frames*.



relate a pair of *HypeStates* $\mathsf{s}$ and $\mathsf{s}'$ to a unique third *HypeState*.[16] But we shall leave further details for some other occasion and continue with our total fusion operation $\oplus$.

Moreover, it should be remembered that Leitgeb explicit labels the partiality of $\circ$ a 'design choice' and concludes with the hope that the "success of the system as a whole is going to justify the design choice" (2019, 323 footnote 9). I similarly suggest that the success of our metaphysical analysis of the two approaches to Routley star and analysis of HYPE has to be judged by the success of what the system can represent as a whole. So let me further suggest that the loss of partiality in the analysis of $\circ$ by $\oplus$ isn't serious because the system of OT as a whole reconstructs the partiality to which $\circ$ is put to use in HYPE by other means. For example, in Example 3 (2019, 335), Leitgeb makes use of the partiality of $\circ$ to build a countermodel in HYPE to the inference $\varphi \rightarrow \bot \models \neg\varphi$. In that countermodel, there are two HYPE states $s_a$ and $s_b$, and though $\circ$ is idempotent, both $s_a \circ s_b$ and $s_b \circ s_a$ are undefined. Leitgeb then shows this model yields a counterexample to $A \rightarrow \bot \models \neg A$.[17]

By contrast, in OT, we don't need $\oplus$ to be partial to construct a *HypeState* $\mathsf{s}$ and a *Hype*-proposition $\mathsf{p}$ such that $\mathsf{s} \models (\mathsf{p} \rightarrow \bot)$ but $\mathsf{s} \not\models \neg\mathsf{p}$. To see this, just take on board the assumption made in HYPE's language that $\top$ (or $\neg\top$ or $\bot$) is a distinguished proposition term. Then let $\mathsf{p}_2$ be an arbitrary *Hype*-proposition. Note that OT doesn't require the identity $\mathsf{p}_2 \rightarrow \bot = \neg\mathsf{p}_2$. So consider the *HypeState* $\mathsf{s}_2$ that makes just $p_2 \rightarrow \bot$ true:

$$\mathsf{s}_2 =_{df} \imath \mathsf{s} \forall \mathsf{p}\big(\mathsf{s} \models \mathsf{p} \equiv \mathsf{p} = (\mathsf{p}_2 \rightarrow \bot)\big)$$

One cannot validly infer in OT that $\mathsf{s}_2 \models \neg\mathsf{p}_2$. And if given the hyperintensional claim that $\mathsf{p}_2 \rightarrow \bot \neq \neg\mathsf{p}_2$, it does follow that $\neg\mathsf{s}_2 \models \neg\mathsf{p}_2$. So,

[16] Intuitively, $R$ would be a partial relation that is idempotent and commutative when $\imath \mathsf{s}_0 R\mathsf{s}\mathsf{s}'\mathsf{s}_0$ exists. Then we could re-define $\oplus$ for *HypeStates* so that it meets the following condition:

$$\mathsf{s} \oplus \mathsf{s}' =_{df} \imath \mathsf{s}'' \forall \mathsf{p}(\mathsf{s}'' \models \mathsf{p} \ \equiv\ \mathsf{s} \models \mathsf{p} \vee \mathsf{s}' \models \mathsf{p} \vee \imath \mathsf{s}_0(R\mathsf{s}\mathsf{s}'\mathsf{s}_0) \models \mathsf{p})$$

The intuition here is that $R$ ensures that $\mathsf{s} \oplus \mathsf{s}'$ makes true *Hype*-propositions other than the ones true in $s$ and $\mathsf{s}'$. Moreover, we must also require:

$$\imath \mathsf{s}_0 R\mathsf{s}_1((\mathsf{s}_1 \oplus \mathsf{s}_2) \oplus \mathsf{s}_3)\mathsf{s}_0 \trianglelefteq (\mathsf{s}_1 \oplus \mathsf{s}_2) \oplus \mathsf{s}_3$$

The extra constraint on $R$ guarantees partial associativity. Thus, constraints on $R$ validate idempotence, commutativity when defined, and partial associativity when defined.

[17] For those familiar with HYPE, the argument is this (mostly quoting from 2019, 335): [It holds that] $s_a \models p_2 \rightarrow \bot$, as there is no $p_2$ state with which $s_a$ could be fused. However, $s_a \not\models \neg p_2$, since $\overline{p_2}$ has not been assigned to $s_a$, or, equivalently (by Lemma 8), because $s_a$ does not stand in the $\bot_3$ relation to the $p_2$-satisfying state $s_b$.

we don't need partiality in $\oplus$ to build a countermodel of the inference $\varphi \to \bot \models \neg\varphi$; instead, we just exploit the hyperintensionality already built into OT's theory of propositions. But if one is, nevertheless, still convinced that the metaphysics of HYPE can't be reconstructed with the partiality of $\circ$, then see footnote 16 for a means of doing so.

### 4.3 The HYPE Explicit Incompatibility Relation

Next we define, in object-theoretic terms, the *explicit incompatibility* condition $\bot$ (not to be confused with the proposition $\bot$ just discussed) that holds between *HypeStates*, by first defining it on situations generally. We say $s$ is *explicitly incompatible* with $s'$ ($s\,!\,s'$) just in case there is a proposition $p$ such that $s$ makes $p$ true and $s'$ makes the negation of $p$ true:

$$s\,!\,s' \equiv_{df} \exists p(s \models p \;\&\; s' \models \overline{p}) \tag{42}$$

Since explicit incompatibility is now defined for all situations, it is defined on *HypeStates*, i.e., we may henceforth write $\mathrm{s}\,!\,\mathrm{s}'$ when *HypeStates* are explicitly incompatible.

Now the first principle governing $\bot$ in HYPE is (Leitgeb 2019, 322):

- If there is a $v$ with $v \in V(s)$ and $\overline{v} \in V(s')$, then $s \bot s'$.

Given our interpretation of $\bot$ in terms of !, this becomes represented and derived as the following theorem governing *HypeStates* and *Hype*-propositions:

$$\vdash (\mathrm{s} \models \mathrm{p} \;\&\; \mathrm{s}' \models \overline{\mathrm{p}}) \to \mathrm{s}\,!\,\mathrm{s}' \tag{43}$$

And the second principle governing $\bot$ in HYPE is (Leitgeb 2019, 322):

- If $s \bot s'$ and both $s \circ s''$ and $s' \circ s'''$ are defined, then $s \circ s'' \bot s' \circ s'''$.

Given our interpretation of $\circ$ as $\oplus$ and the fact that $s \oplus s'$ is always defined for any situations $s$ and $s'$, this becomes represented and derived as the following theorem regarding *HypeStates*:

$$\vdash \mathrm{s}\,!\,\mathrm{s}' \to (\mathrm{s} \oplus \mathrm{s}'')\,!\,(\mathrm{s}' \oplus \mathrm{s}''') \tag{44}$$

The proofs of both (43) and (44) are in the Appendix.[18]

[18]In addition to analyzing the HYPE incompatibility condition $\bot$ in terms of the object-theoretic definition of ! in (42), we may also analyze the compatibility relation $C$ used in Restall 2000 (853), Berto 2015 (767), and Berto & Restall 2019 (1127) as the negation of !, i.e., via the following definition, cast in terms of situations generally:

$$sCs' \equiv_{df} \neg\exists p(s \models p \;\&\; s' \models \overline{p})$$

That is, $s$ is compatible with $s'$ just in case there is no proposition $p$ that $s$ makes $p$ true and $s'$ makes $\overline{p}$ true. (Depending on the purpose at hand, one might prefer to revise this definition to ensure that compatibility holds only when no proposition $p$ is such that $p$ is true in the modal closure of $s$ while $\overline{p}$ is true in the modal closure of $s'$. But we don't need this more sophisticated understanding of compatibility in what follows.) Then, the semantic principle governing $C$ stipulated in Restall 2000 (853, Definition 1.1), namely:

for any $x$, $y$, $x'$, and $y'$, if $xCy$, $x' \sqsubseteq x$, and $y' \sqsubseteq y$, then $x'Cy'$,

becomes derivable in OT, with $\trianglelefteq$ instead of $\sqsubseteq$. Let us temporarily use $x,y$ as restricted variables ranging over situations. Then we have:

$$\vdash (xCy \;\&\; x' \trianglelefteq x \;\&\; y' \trianglelefteq y) \to x'Cy'$$

*Proof*: Assume $xCy$, $x' \trianglelefteq x$, and $y' \trianglelefteq y$. These assumptions imply, respectively:

(A) $\neg\exists p(x \models p \;\&\; y \models \overline{p})$

(B) $\forall p(x' \models p \to x \models p)$

(C) $\forall p(y' \models p \to y \models p)$

To show $x'Cy'$, suppose not, for reductio. Then $\exists p(x' \models p \;\&\; y' \models \overline{p})$. Suppose $p_1$ is such a proposition, so that we know both $x' \models p_1$ and $y' \models \overline{p_1}$. Then by (B) and (C), respectively, these entail that $x \models p_1$ and $y \models \overline{p_1}$. But this contradicts (A).

As noted in footnote 3, this same principle, labeled 'Backwards' (compatibility), is stipulated in Berto 2015 (768) and Berto & Restall 2019 (1129).



## 5 Routley Star in HYPE

We continue to use our restricted variables 'p' and 's' to range over *Hype*-propositions and *HypeStates*, respectively. Our next goal, then, is to reconstruct and derive the principles that govern the HYPE Routley star operator. That is, we must reconstruct and derive the following conditions laid down in Leitgeb 2019 (322), in which we've replaced Leitgeb's variable '$s$' by our restricted variable 's':

For every $\mathrm{s}$ in $S$,

(A) there is a unique $\mathrm{s}^* \in S$ (the star image of s) such that:

(B) $V(\mathrm{s}^*) = \{\overline{v} \mid v \notin V(\mathrm{s})\}$,

(C) $\mathrm{s}^{**} = \mathrm{s}$,

(D) s and $\mathrm{s}^*$ are not incompatible, i.e., $\neg(\mathrm{s} \bot \mathrm{s}^*)$, and

(E) $\mathrm{s}^*$ is the largest state compatible with s, i.e., if s is not incompatible with $\mathrm{s}'$, then the fusion of $\mathrm{s}'$ and $\mathrm{s}^*$ is defined and the fusion of $\mathrm{s}' \circ \mathrm{s}^* = \mathrm{s}^*$.

Note that $s^*$ is defined in HYPE as $V(s^*) = \{\overline{v} | v \notin V(s)\}$, instead of as $V(s^*) = \{v | \overline{v} \notin V(s)\}$. However, as we saw in Section 3, these two definitions become equivalent if propositions and their double negations are generally identified. And as we saw in Section 4, Leitgeb does identify p and $\overline{\overline{p}}$ in his propositional language $\mathcal{L}$. Since we've defined *Hype*-propositions as ones that exhibit this behavior, let us examine how the HYPE Routley star and the principles governing it can be defined or derived given our analysis of *Hype*-propositions and *HypeStates*.

For any *HypeState* s, we may define the HYPE Routley star image of s, written $s^*$, as the *HypeState* $s'$ that makes a *Hype*-proposition p true just in case p is the negation of a proposition not true in s:[19]

$$s^* =_{df} \imath s' \forall p(s' \models p \equiv \exists q(\neg s \models q \,\&\, p = \overline{q})) \tag{45}$$

We take (45) to be a reconstruction of principle (B) above. Now although the HYPE principle (A) requires that there be a unique $s^*$ satisfying (B) – (E), it should be clear that $s^*$ is already uniquely defined; for any s, exactly one $s^*$ has been identified by a canonical description.

So we may immediately conclude that $s^*$ exists, for any s. Before we show that s's unique star image $s^*$ also satisfies constraints (C) – (E), it proves useful to first confirm a few facts that follow from (45).

By now familiar reasoning, we may infer from (45) that the *Hype*-propositions true in $s^*$ are precisely the negations of the *Hype*-propositions that fail to be true in s:

$$\vdash \forall p(s^* \models p \equiv \exists q(\neg s \models q \,\&\, p = \overline{q})) \tag{46}$$

Moreover, we may verify that the principle proved in Section 3 holds for *HypeStates*, namely that p is the negation of some proposition that s fails to make true if and only if s fails to make $\overline{p}$ true:

$$\vdash \exists q(\neg s \models q \,\&\, p = \overline{q}) \equiv \neg s \models \overline{p} \tag{47}$$

Clearly, then, (46) and (47) imply that $s^*$ makes p true if and only if s fails to make $\overline{p}$ true; and by simple logical consequence of this fact, it follows that $\overline{p}$ is true in a *HypeState* s if and only if it is not the case that p is true in $s^*$:

[19]The following should be considered a *redefinition* of the Routley star image. That's because *HypeStates* are situations and, in Section 2, (15) defines the Routley star on situations. So to avoid conflicting definitions, just consider the following as a redefinition of this operator.



$$\vdash \forall p(s^* \models p \equiv \neg s \models \overline{p}) \tag{48}$$

$$\vdash \forall p(s \models \overline{p} \equiv \neg s^* \models p) \tag{49}$$

(48) is a direct analogue of the Routley & Routley condition (iv) described in the Introduction above, and so corresponds directly to (16).

Note next that we can make use of the definitions of gaps and gluts in (18) and (19), respectively; these notions were defined generally for any situations and propositions and so apply to *HypeStates* and *Hype*-propositions. We may then further confirm that (45) is correct by establishing that if s has a glut w.r.t. p, then $s^*$ has a gap w.r.t. p; if s has a gap w.r.t. p, then $s^*$ has a glut w.r.t. p; and if s has neither a glut nor a gap w.r.t. p, then $s^*$ agrees with $s^*$ on p:

$$\vdash \mathit{GlutOn}(s,p) \to \mathit{GapOn}(s^*,p) \tag{50}$$

$$\vdash \mathit{GapOn}(s,p) \to \mathit{GlutOn}(s^*,p) \tag{51}$$

$$\vdash (\neg \mathit{GlutOn}(s,p) \,\&\, \neg \mathit{GapOn}(s,p)) \to (s^* \models p \equiv s \models p) \tag{52}$$

Now that we have confirmed that (45) is a definition of $s^*$ that yields the latter's desired characteristics, we turn to the derivation of principle (C) governing HYPE $s^*$, namely, that $s^{**}$ is identical to s:

$$\vdash s^{**} = s \tag{53}$$

Cf. Leitgeb 2019 (322). So, whereas (29) establishes that the stipulation $s^{**} = s$ in Routley & Routley 1973 is equivalent to the double-negation condition $\forall p(s \models p \equiv s \models \overline{\overline{p}})$, (53) establishes that the analogous stipulation $s^{**} = s$ in Leitgeb 2019 can be derived from the double-negation fact about *Hype*-propostions that $\forall p(p = \overline{\overline{p}})$. These results give us a deeper understanding of the connection between the two ways of defining the Routley star image of a situation.

Principles (D) and (E) of HYPE $s^*$ may be derived as follows. (D) asserts that s is not explicitly incompatible with $s^*$:

$$\vdash \neg s!s^* \tag{54}$$

And since $s' \oplus s^*$ is always defined in our reconstruction, we can reconstruct and derive (E) as the simpler claim if s is not incompatible with $s'$, then the sum/fusion of $s'$ and $s^*$ just is $s^*$:

$$\vdash \neg s!s' \to (s' \oplus s^* = s^*) \tag{55}$$

($55$) guarantees that $\mathrm{s}^*$ is the largest state compatible with s.

Finally, if we recall the definition $s \trianglelefteq s'$ (39) and the fact that $s \trianglelefteq s' \equiv \forall p(s \models p \rightarrow s' \models p)$ (40), we may prove that the Routley star operation reverses $\trianglelefteq$:

$$\vdash \mathrm{s} \trianglelefteq \mathrm{s}' \rightarrow \mathrm{s}'^* \trianglelefteq \mathrm{s}^* \tag{56}$$

Cf. Observation 3, Leitgeb 2019 (325). This completes the derivation of the principles stipulated in HYPE for the Routley star operation, modulo the partiality of the HYPE fusion operation.

## 6 Conclusion

We've now answered the question: What kind of metaphysics is represented by a semantics making use of Routley star? Without assuming any mathematical entities or theory of sets and functions, we've used OT to define two forms of the Routley star operation and derive the principles that govern these forms. And the better we understand the theorems that are implied by the two ways of defining it, the better we understand how the star operation might be used. The existence of the Routley star image $s^*$ of a situation $s$ is guaranteed not by set theory but by a theory of abstract objects. And out reconstruction shows that situations have both a metaphysical character and an informational character, at least as these are described in the quote above from Leitgeb 2019 (footnote 9). One can view situations in OT as "chunks of reality" that are "located in the world", especially if one takes an Aristotelian view of abstract objects as forms that are part of reality. Alternatively, one can view situations informationally, as abstract entities systematize the content of information that might be communicated. But these metaphilosophical considerations about how to interpret OT as a theory shouldn't divert attention away from the tight conceptual framework that OT provides for defining Routley star.

Indeed, if you look at how situations are defined in (2) and at how the Routley star operation is defined in (15) and (45), one might even suggest that the star operation is a logical one. Propositions are axiomatized as 0-ary relations and can be considered part of logic. Situations are defined in (2) as abstract objects that encode only propositional properties. And the $^*$ operation is then defined on situations in terms of the notions *the*, *truth in* (which is in turn defined in terms of the *encoding* mode of predication), *every* and *some*, *if and only if*, and *not*. If the star operation is logical, then we can explain why some have thought that the uses to which Routley star has been put in the literature helps us to capture semantically a more general and flexible logical concept of negation.[20]

Finally, we've shown that the basic principles governing Routley star need not be stipulated but can be derived from its definition. This integrates Routley star into a more general theory of (partial) situations that has been shown, in previous work, to ground the theory of both possible worlds and impossible worlds. This analysis of the Routley star operation clarifies our understanding of the Routley-Meyer ternary relation $R$ (Routley-Meyer 1972, 1973) on 'set-ups', by systematically validating many of the assumptions of situation theory used in Mares' (2004) motivation and justification for $R$. But we shall not attempt to further explore the various definitions of the ternary relation $R$ in this paper. It is sufficient to have shown how different groups of situations (e.g., those defined in (2), *HypeStates*, or others) can constitute a proper subdomain for ternary $R$. I take those subdomains to be consistent with all of the various attempts at understanding that relation.

## Appendix: Proofs of the Theorems

**(8)**[21] If we eliminate the restricted variable, then the theorem we have to prove becomes:

$$\exists x(Situation(x) \,\&\, \forall p(x \models p \equiv \varphi)), \text{ provided } x \text{ isn't free in } \varphi$$

So let $\varphi$ be any formula in which $x$ doesn't occur free. (Note that the variable $p$ may or may not be free in $\varphi$.) Now, pick a property variable that doesn't occur free in $\varphi$. Without loss of generality, suppose it is $G$. Then let $\psi$ be the formula $\exists p(\varphi \& G = [\lambda z p])$. Clearly, since $x$ doesn't occur free in $\psi$, the following is a schematic instance of of the comprehension schema for abstract objects, stated in the text as (1):

$$\exists x(A!x \,\&\, \forall G(xG \equiv \psi))$$

But given our choice of $\psi$, this amounts to:

$$\exists x\big(A!x \,\&\, \forall G(xG \equiv \exists p(\varphi \,\&\, G = [\lambda z\, p]))\big)$$

[20] I'm indebted to Hannes Leitgeb for suggesting this point.

[21] I'm indebted to Uri Nodelman for spotting a flaw in the original proof of this theorem.

Let $a$ be such an object, so that we know both $A!a$ and:

(A) $\forall G(aG \equiv \exists p(\varphi \,\&\, G = [\lambda z\, p]))$

It follow *a fortiori* that $\forall G(aG \rightarrow \exists p(G = [\lambda z\, p]))$. Hence *Situation*$(a)$, by definition (2). So it remains to show $\forall p(a \models p \equiv \varphi)$. By GEN, it suffices to show $a \models p \equiv \varphi$, since we've made no special assumptions about $p$.

To prove this biconditional, we'll rely on the fact that $a \models p$ is defined as $a[\lambda z\, p]$, by (3), given that $a$ is a situation. We'll therefore want to instantiate $a[\lambda z\, p]$ into (A). But there is a clash of variables and, to avoid this, we use the following alphabetic variant of (A), where $q$ is a variable that is substitutable for $p$, and doesn't occur free, in $\varphi$:

(A′) $\forall G(aG \equiv \exists q(\varphi^q_p \,\&\, G = [\lambda z\, q]))$

Now we can properly instantiate $[\lambda z\ p]$ into $(\theta')$, and if we remember that $G$ doesn't occur free in $\varphi$, we obtain:[22]

(B) $a[\lambda z\, p] \equiv \exists q(\varphi^q_p \,\&\, [\lambda z\, p] = [\lambda z\, q])$

With these facts we can prove $a \models p \equiv \varphi$.

$(\rightarrow)$ Assume $a \models p$, to show $\varphi$. Then $a[\lambda z\, p]$, by (3). So by (B), it follows that:

$$\exists q(\varphi^q_p \,\&\, [\lambda z\, p] = [\lambda z\, q])$$

Now suppose $q_1$ is such a proposition, so that we know:

(C) $(\varphi^q_p)^{q_1}_q \,\&\, [\lambda z\, p] = [\lambda z\, q_1]$

In OT, propositions are identical whenever the propositional properties constructed from them are identical (Zalta 1993, 409). So by the second conjunct of (C), it follows that $p = q_1$. Hence, by the first conjunct of (C), it follows that $(\varphi^q_p)^p_q$. But since the conditions of the Re-replacement Lemma are met (Enderton 2001, 130), this latter is just $\varphi$.

$(\leftarrow)$ Assume $\varphi$. Then $\varphi \,\&\, [\lambda z\, p] = [\lambda z\, p]$, by the reflexivity of identity. Hence, by existential introduction:

$$\exists q(\varphi^q_p \,\&\, [\lambda z\, p] = [\lambda z\, q])$$

[22] Strictly speaking, when we instantiate $[\lambda z\, p]$ into (A′), we obtain:

$$a[\lambda z\, p] \equiv \exists q((\varphi^q_p)^{[\lambda z\, p]}_G \,\&\, [\lambda z\, p] = [\lambda z\, q])$$

But since $G$ isn't free in $\varphi$, $(\varphi^q_p)^{[\lambda z\, p]}_G$ is just $\varphi^q_p$.



Then by (B), $a[\lambda z\, p]$. So by (3) and the fact that $a$ is a situation, $a \models p$. ⋈

**(9)** This is Theorem 2 in Zalta 1993. The proof was given in Zalta 1991 (Appendix A), which served as a precursor to Zalta 1993.

**(10)** This follows from (8) and (9) by the standard definition of the uniqueness quantifier $\exists! s\psi$.

**(12)** Suppose $y$ is substitutable for $x$ in $\varphi$ and assume $y = \iota x\varphi$. Then by axiom (11), $\forall x(\mathcal{A}\varphi \equiv x = y)$. But since $y$ is substitutable for $x$ in $\varphi$, we can instantiate this last fact to $y$ and we obtain $\mathcal{A}\varphi^y_x \equiv y = y$. So by the reflexivity of identity, $\mathcal{A}\varphi^y_x$. ⋈

**(13)** By hypothesis, $\varphi$ is modally collapsed and $y$ is substitutable for $x$ in $\varphi$. Now assume $y = \iota x\varphi$, to show $\varphi^y_x$. It follows from this assumption by theorem (12) that $\mathcal{A}\varphi^y_x$. But since $\varphi$ is modally collapsed, there is a proof of $\Box(\varphi \rightarrow \Box\varphi)$. Since this latter is a theorem, it follows by GEN that $\forall x\Box(\varphi \rightarrow \Box\varphi)$. Instantianting to $y$ it follows that $\Box(\varphi^y_x \rightarrow \Box\varphi^y_x)$. But as we saw in footnote 11, a formula of this form implies $\mathcal{A}\varphi^y_x \equiv \varphi^y_x$. Hence, $\varphi^y_x$. ⋈

**(14)** Suppose $s'$ isn't free in $\varphi$ and $\varphi$ is modally collapsed. To show:

$$\big(s = \iota s'\forall p(s' \models p \equiv \varphi)\big) \rightarrow \forall p(s \models p \equiv \varphi)$$

it suffices to show that the formula $\forall p(s' \models p \equiv \varphi)$ is modally collapsed, for then our theorem becomes an instance of (13). So we have to prove:

$$\Box\big(\forall p(s' \models p \equiv \varphi) \rightarrow \Box\forall p(s' \models p \equiv \varphi)\big)$$

By the Rule of Necessitation, it suffices to prove:

$$\forall p(s' \models p \equiv \varphi) \rightarrow \Box\forall p(s' \models p \equiv \varphi)$$

So assume $\forall p(s' \models p \equiv \varphi)$, to show $\Box\forall p(s' \models p \equiv \varphi)$. By the Barcan Formula, it suffices to show $\forall p\Box(s' \models p \equiv p)$. Since $p$ isn't free in our assumption, it remains, by GEN, to show $\Box(s' \models p \equiv p)$. So $p$ is a fixed, but arbitrary proposition, and so our assumption that $\forall p(s' \models p \equiv \varphi)$ implies:

(A) $s' \models p \equiv \varphi$

By hypothesis, $\varphi$ is modally collapsed, and so we know that the following is a theorem:

(B) $\Box(\varphi \rightarrow \Box\varphi)$

But independently, note that $s' \models p$ is defined in (3) as $s'[\lambda y\, p]$, and so it is a formula of the form $xF$. Since the modal logic of encoding is expressed by the principle $xF \rightarrow \Box xF$ (Zalta 1993, 403), it follows by the Rule of Necessitation that $\Box(xF \rightarrow \Box xF)$. Hence as an instance, we know:

(C) $\Box(s' \models p \rightarrow \Box s' \models p)$

But it is a theorem of modal logic that if formulas $\psi$ and $\chi$ necessarily imply their own necessity, then the material equivalence of $\psi$ and $\chi$ necessarily implies their necessary equivalence:

$$(\Box(\psi \rightarrow \Box\psi)\ \&\ \Box(\chi \rightarrow \Box\chi)) \rightarrow \Box((\psi \equiv \chi) \rightarrow \Box(\psi \equiv \chi))$$

Given this theorem and setting $\psi$ to $s \models p$ and $\chi$ to $\varphi$, (C) and (B) jointly imply:

$$\Box((s' \models p \equiv \varphi) \rightarrow \Box(s' \models p \equiv \varphi))$$

So by the T schema,

$$(s' \models p \equiv \varphi) \rightarrow \Box(s' \models p \equiv \varphi)$$

Hence, by (A), $\Box(s' \models p \equiv \varphi)$, which is what it remained to show. ⋈

**(16)** First, we show that $\neg s \models \overline{p}$ is a modally collapsed formula:

*Lemma*: $\Box(\neg s \models \overline{p} \rightarrow \Box\neg s \models \overline{p})$

*Proof*. By the Rule of Necessitation, it suffices to prove $\neg s \models \overline{p} \rightarrow \Box\neg s \models \overline{p}$. So assume $\neg s \models \overline{p}$, to show $\Box\neg s \models \overline{p}$. Now, as previously noted in the text, the modal logic of encoding is $xF \rightarrow \Box xF$. So, by the T schema and the Rule of Necessitation, we know $\Box(xF \equiv \Box xF)$. This implies $\Box(\Diamond xF \equiv xF)$. As an instance of this latter, $\Box(\Diamond s \models \overline{p} \equiv s \models \overline{p})$. Then by the T schema, $\Diamond s \models \overline{p} \equiv s \models \overline{p}$. So, negating both sides, $\neg\Diamond s \models \overline{p} \equiv \neg s \models \overline{p}$. Then by our assumption, it follows that $\neg\Diamond s \models \overline{p}$, which is equivalent to $\Box\neg s \models \overline{p}$, which is what we had to show.

Now note that we can apply GEN to (14), since $s$ is a free variable, to conclude:

$$\forall s\big(s = \imath s'\forall p(s' \models p \equiv \varphi)\big) \rightarrow \forall p(s \models p \equiv \varphi),$$
provided $s'$ isn't free in $\varphi$ and $\varphi$ is modally collapsed



Now since $s'$ isn't free in $\neg s \models \overline{p}$ and this formula is modally collapsed, we can let $\varphi$ be $\neg s \models \overline{p}$, so that as an instance of the foregoing, we know:

$$\forall s\big(s = \imath s'\forall p(s' \models p \equiv \varphi)\big) \rightarrow \forall p(s \models p \equiv \neg s \models \overline{p})$$

So we may instantiate $s^*$ into this universal claim and the result is:

$$s^* = \imath s'\forall p(s' \models p \equiv \neg s \models \overline{p}) \rightarrow \forall p(s^* \models p \equiv \neg s \models \overline{p})$$

So by definition (15), $\forall p(s^* \models p \equiv \neg s \models \overline{p})$. ⋈

**(17)** By (16) we know:

(A) $\forall p(s^* \models p \equiv \neg s \models \overline{p})$

Since $\varphi \equiv \neg\psi$ is necessarily equivalent to $\neg\varphi \equiv \psi$, it follows from (A) by the Rule of Substitution that:

(B) $\forall p(\neg s^* \models p \equiv s \models \overline{p})$

And since $\varphi \equiv \psi$ is necessarily equivalent to $\psi \equiv \varphi$, it follows from (B) by the Rule of Substitution that:

$$\forall p(s \models \overline{p} \equiv \neg s^* \models p)$$ ⋈

**(20)** Take the following as a global assumption:

(A) $s = s^{**}$

We want to prove that if *GlutOn*$(s,p)$, then *GapOn*$(s^*,p)$. So assume *GlutOn*$(s,p)$, i.e., by (18), that:

(B) $s \models p\ \&\ s \models \overline{p}$

To show *GapOn*$(s^*,p)$, we have to show both (a) $\neg s^* \models p$ and (b) $\neg s^* \models \overline{p}$, by (19).

(a) If we instantiate (17) to $s$ and $p$, we obtain:

$$s \models \overline{p} \equiv \neg s^* \models p$$

So by the 2nd conjunct of (B), $\neg s^* \models p$.

(b) If we instantiate (17) to $s^*$ and $p$, we obtain:

(C) $s^* \models \overline{p}\ \equiv\ \neg s^{**} \models p$

But the 1st conjunct of (B) implies, under our global assumption $s = s^{**}$ (A), that $s^{**} \models p$. But this fact and (C) jointly imply $\neg s^* \models \overline{p}$. ⋈

**(21)** Take the following as a global assumption:

(A) $s = s^{**}$

We want to prove that if $GapOn(s,p)$, then $GlutOn(s^*,p)$. So assume $GapOn(s,p)$, i.e., by (19), that:

(B) $\neg s \models p \,\&\, \neg s \models \overline{p}$

Then to show $GlutOn(s^*,p)$, we show both (a) $s^* \models p$ and (b) $s^* \models \overline{p}$, by (18).

(a) If we instantiate (16) to $s$ and $p$, we obtain:

$$s^* \models p \equiv \neg s \models \overline{p}$$

This result and the second conjunct of (B) imply $s^* \models p$.

(b) If we instantiate (17) to $s^*$ and $p$, we obtain:

(C) $s^* \models \overline{p} \equiv \neg s^{**} \models p$

But given our global assumption (A) that $s = s^{**}$, it follows from the first conjunct of (B) that $\neg s^{**} \models p$. But from this fact and (C), it follows that $s^* \models \overline{p}$. ⋈

**(22)** Assume both $\neg GlutOn(s,p)$ and $\neg GapOn(s,p)$. Then by definitions (18) and (19), we know:

$$\neg(s \models p \,\&\, s \models \overline{p})$$

$$\neg(\neg s \models p \,\&\, \neg s \models \overline{p})$$

These are, respectively, equivalent to:

(A) $\neg s \models p \vee \neg s \models \overline{p}$

(B) $s \models p \vee s \models \overline{p}$

We may then prove both directions of $s^* \models p \equiv s \models p$. (→) Assume $s^* \models p$. Then by (16), $\neg s \models \overline{p}$. It follows from this and (B) that $s \models p$. (←) Assume $s \models p$. This and (A) imply $\neg s \models \overline{p}$. So by (16), $s^* \models p$. ⋈

**(23)** Assume:

$$\forall p(\neg GlutOn(s,p) \,\&\, \neg GapOn(s,p))$$

To show $s^* = s$, we have to show $\forall p(s^* \models p \equiv \neg s \models p)$, by (9). By GEN, we show $s^* \models p \equiv s \models p$. But if we instantiate our assumption to $p$, we obtain $\neg GlutOn(s,p) \,\&\, \neg GapOn(s,p)$, and so $s^* \models p \equiv s \models p$ follows by (22). ⋈



**(24)** (→) Our (global) assumption is:

$$\forall p(\neg GlutOn(s,p) \,\&\, \neg GapOn(s,p))$$

We want to show $\forall p(s \models p \equiv \neg s \models \overline{p})$. By GEN, it suffices to show $s \models p \equiv \neg s \models \overline{p}$. But it is an immediate consequence of our global assumption that:

(A) $\neg GlutOn(s,p) \,\&\, \neg GapOn(s,p)$

We use this to prove both directions of our biconditional:

(→) Assume (locally) $s \models p$. The first conjunct of (A) and definition (18) imply $\neg(s \models p \,\&\, s \models \overline{p})$, i.e., $\neg s \models p \vee \neg s \models \overline{p}$. This last fact and our local assumption jointly imply $\neg s \models \overline{p}$.

(←) Assume (locally) $\neg s \models \overline{p}$. The second conjunct of (A) and definition (19) imply $\neg(\neg s \models p \,\&\, \neg s \models \overline{p})$, i.e., $s \models p \vee s \models \overline{p}$. But his last fact and our local assumption jointly imply $s \models p$.

(←) Our (global) assumption is:

$$\forall p(s \models p \equiv \neg s \models \overline{p})$$

To show $\forall p(\neg GlutOn(s,p) \,\&\, \neg GapOn(s,p))$, it suffices by &I and GEN and to show both (a) $\neg GlutOn(s,p)$ and (b) $\neg GapOn(s,p)$. But it is an immediate consequence of our global assumption that:

(B) $s \models p \equiv \neg s \models \overline{p}$

We use this to show both directions of our biconditional:

(→) Assume, for reductio, that $GlutOn(s,p)$. Then by definition (18), we know both $s \models p$ and $s \models \overline{p}$. But the former implies the negation of the latter, by (B). Contradiction.

(←) Assume, for reductio, that $GapOn(s,p)$. Then by (19), we know both $\neg s \models p$ and $\neg s \models \overline{p}$. But again, the former implies the negation of the latter, by (B). Contradiction. ⋈

**(27)** Take as our global assumption that $\forall s(s^{**} = s)$. From definition (25) and the fact that the condition $p \neq p$ is modally collapsed (by the necessity of identity), it follows that $\forall p(\boldsymbol{s}_{\emptyset} \models p \equiv p \neq p)$, by (14). But since no proposition fails to be self-identical, it follows from this last fact that $\neg\exists p(\boldsymbol{s}_{\emptyset} \models p)$. This implies $\forall p \neg(\boldsymbol{s}_{\emptyset} \models p)$. Now let $q$ be an arbitrarily chosen proposition, so that we know both $\neg \boldsymbol{s}_{\emptyset} \models q$ and $\neg \boldsymbol{s}_{\emptyset} \models \overline{q}$. Then by definition (19), $GapOn(\boldsymbol{s}_{\emptyset}, q)$. But given our global assumption, we know ${s_{\emptyset}}^{**} = s_{\emptyset}$. So by the relevant instance of (21), it follows from $GapOn(\boldsymbol{s}_{\emptyset}, q)$ that $GlutOn({\boldsymbol{s}_{\emptyset}}^*, q)$. From this, it follows *a fortiori* by definition (18) that ${\boldsymbol{s}_{\emptyset}}^* \models q$. Since $q$ was arbitrary, we have established:

(A) $\forall p({\boldsymbol{s}_{\emptyset}}^* \models p)$

But, independently, we also know, given definition (26) and the fact that the condition $p = p$ is modally collapsed (by the necessity of identity), that $\forall p(\boldsymbol{s}_{\boldsymbol{V}} \models p \equiv p = p)$. Since every proposition is self-identical, it follows from this last fact that:

(B) $\forall p(\boldsymbol{s}_{\boldsymbol{V}} \models p)$

Now $\forall p\varphi \;\&\; \forall p\psi$ implies $\forall p(\varphi \equiv \psi)$. So we may conclude from (A) and (B) that:

$$\forall p({\boldsymbol{s}_{\emptyset}}^* \models p \equiv \boldsymbol{s}_{\boldsymbol{V}} \models p)$$

Since ${\boldsymbol{s}_{\emptyset}}^*$ and $\boldsymbol{s}_{\boldsymbol{V}}$ are situations that make the same propositions true, it follows by (9) that ${\boldsymbol{s}_{\emptyset}}^* = \boldsymbol{s}_{\boldsymbol{V}}$. ⋈

**(28)** (Exercise)

**(29)** ($\rightarrow$) Assume $s^{**} = s$. By GEN, it suffices to show $s \models p \equiv s \models \overline{\overline{p}}$. The identity of $s^{**}$ and $s$ implies, by (9), that $\forall p(s^{**} \models p \equiv s \models p)$. Hence $s^{**} \models p \equiv s \models p$, which commutes to:

(A) $s \models p \equiv s^{**} \models p$

Now, independently, if we instantiate (16) to $s^*$ and $p$, we also know:

(B) $s^{**} \models p \equiv \neg s^* \models \overline{p}$

Moreover, independently, we know $s^* \models \overline{p} \equiv \neg s \models \overline{\overline{p}}$, by instantiating (16) to $s$ and $\overline{p}$. By negating both sides and eliminating the double negation, we have:

(C) $\neg s^* \models \overline{p} \equiv s \models \overline{\overline{p}}$

So $s \models p \equiv s \models \overline{\overline{p}}$, by biconditional syllogism from (A), (B), and (C).

($\leftarrow$) Assume:

(D) $\forall p(s \models p \equiv s \models \overline{\overline{p}})$

To establish $s^{**} = s$, we appeal to (9) and show $\forall p(s^{**} \models p \equiv s \models p)$. By GEN, it suffices to show $s^{**} \models p \equiv s \models p$. First note that, by GEN, (16) holds for all $s$ and so if we instantiate the resulting universal claim to $s^*$ and $p$, we obtain:

(E) $s^{**} \models p \equiv \neg s^* \models \overline{p}$

Independently, we obtain $s^* \models \overline{p} \equiv \neg s \models \overline{\overline{p}}$ by instantiating (16) to $s$ and $\overline{p}$. This is equivalent to:

(F) $\neg s^* \models \overline{p} \equiv s \models \overline{\overline{p}}$

Moreover, if instantiate (D) to $p$ and commute the result, we know:

(G) $s \models \overline{\overline{p}} \equiv s \models p$

But now, (E), (F), and (G) jointly imply:

$$s^{**} \models p \equiv s \models p \qquad ⋈$$

**(31)** Assume $Hype(p)$. Then by (30), $p = \overline{\overline{p}}$. So we may substitute $\overline{\overline{p}}$ for the first occurrence of $p$ in the identity $\overline{p} = \overline{p}$, to obtain $\overline{\overline{\overline{p}}} = \overline{p}$. So by definition (30), $Hype(\overline{p})$. ⋈

**(34)** By reasoning analogous to (8).

**(35)** By (34) and the definition of identity for situations (9).

**(38)** This is a consequence of (37) and (14), and the fact that $s \models p \vee s' \models p$ is modally collapsed. ⋈

**(40)** We prove both directions.

($\rightarrow$) Assume $s \trianglelefteq s'$. It follows that $\forall p(s \models p \rightarrow s' \models p)$, by definition (39). Now to show $s \oplus s' = s'$, we have to show that $s \oplus s'$ and $s'$ make the same propositions true, by (9). That is, we have to show, for an arbitrary $p$, that $s \oplus s' \models p \equiv s' \models p$. But both directions of this biconditional hold. If $s \oplus s' \models p$ then either $s \models p$ or $s' \models p$, by (38). But in either case, $s' \models p$, given that every proposition true in $s$ is true in $s'$. And if $s' \models p$, then clearly, by a fact about $\oplus$ (38), it follows that $s \oplus s' \models p$.

($\leftarrow$) Assume $s \oplus s' = s'$. It follows by (9) that $s \oplus s'$ and $s'$ make the same propositions true. Now to show $s \trianglelefteq s'$, we need to show, for an arbitrary proposition $p$, that $s \models p \rightarrow s' \models p$. So assume $s \models p$, to show $s' \models p$. But since $s \oplus s'$ and $s'$ make the same propositions true, it suffices to show $s \oplus s' \models p$. But this follows from our assumption that $s \models p$, by (38). ⋈

**(41)** The idempotence, commutativity, and associativity of $\oplus$ with respect to situations and, *a fortiori*, *HypeStates*, follows from (38) and the the facts that $\vee$ is idempotent, commutative, and associative. ⋈

**(43)** This follows from the definition of $s \,!\, s'$ (42) once it is instantiated when to *HypeStates* s and $s'$.

**(44)** Assume $s \,!\, s'$. Then by definition (42), we know $\exists p(s \models p \;\&\; s' \models \overline{p})$. Suppose $p_1$ is such a proposition, so that we know $s \models p_1$ and $s' \models \overline{p_1}$. But since $s \models p_1$, so does $s \oplus s''$, by theorem (38). And by that same theorem, since $s' \models \overline{p_1}$, so does $s' \oplus s'''$. Hence:

$$\exists p\big((s \oplus s'' \models p) \;\&\; (s' \oplus s''' \models \overline{p})\big)$$

So by definition (42), $(s \oplus s'') \,!\, (s' \oplus s''')$. ⋈

**(46)** (Exercise)

**(47)** By reasoning analogous to the proof of ($\omega$) in Section 3, though stated in terms of *Hype*-propositions and *HypeStates*. ⋈

**(48)** This follows from (46) by (47) and the Rule of Substitution. ⋈

**(49)** (Exercise)

**(50)** Assume *GlutOn*(s, p), i.e., by (18) that:

(A) $s \models p$

(B) $s \models \overline{p}$

We want to to show *GapOn*($s^*$, p), i.e., by (19), that both (a) $\neg s^* \models p$ and (b) $\neg s^* \models \overline{p}$. (a) This follows from (B) by (49). (b) If we instantiate (48) to $\overline{p}$, we have $s^* \models \overline{p} \equiv \neg s \models \overline{\overline{p}}$. But this is equivalent to $\neg s^* \models \overline{p} \equiv s \models \overline{\overline{p}}$. Since *Hype*-propositions are identical to their double-negations (30), it follows that $\neg s^* \models \overline{p} \equiv s \models p$. Then by (A), we may infer $\neg s^* \models \overline{p}$. ⋈

**(51)** Assume *GapOn*(s, p), i.e., by (19):



(A) $\neg s \models p$

(B) $\neg s \models \overline{p}$

We want to show *GlutOn*($s^*$, p), i.e., by (18), that both (a) $s^* \models p$ and (b) $s^* \models \overline{p}$. (a) This follows from (B) by (48). (b) If we instantiate (48) to $\overline{p}$, we have $s^* \models \overline{p} \equiv \neg s \models \overline{\overline{p}}$. Since *Hype*-propositions are identical to their double-negations (30), it follows that $s^* \models \overline{p} \equiv \neg s \models p$. From this and (A) it follows that $s^* \models \overline{p}$. ⋈

**(52)** This follows by applying the reasoning in (22) to *HypeStates* and *Hype*-propositions. ⋈

**(53)** To establish $s^{**} = s$, we note that since *HypeStates* encode only *Hype*-propositions (33), it suffices by (9) to show $\forall p(s^{**} \models p \equiv s \models p)$. By GEN, it then suffices to show $s^{**} \models p \equiv s \models p$. Now if we instantiate (48) to $s^*$, we obtain:

(A) $s^{**} \models p \equiv \neg s^* \models \overline{p}$

Independently, if instantiate (49) to $\overline{p}$, we obtain $s \models \overline{\overline{p}} \equiv \neg s^* \models \overline{p}$, which by the commutativity of the biconditional, implies:

$$\neg s^* \models \overline{p} \equiv s \models \overline{\overline{p}}$$

And since *Hype*-propositions are identical with their double negations, it follows from this last result that:

(B) $\neg s^* \models \overline{p} \equiv s \models p$

But (A) and (B) imply $s^{**} \models p \equiv s \models p$. ⋈

**(54)** Assume, for reductio, that $s \,!\, s^*$. So by definition (42), $\exists p(s \models p \;\&\; s^* \models \overline{p})$. Let $q_1$ be such a proposition, so that we know $s \models q_1$ and $s^* \models \overline{q_1}$. By a key fact about $s^*$ (48), the latter implies $\neg s \models \overline{\overline{q_1}}$. But since *Hype*-propositions are identical with their double negations, it follows that $\neg s \models q_1$. Contradiction. ⋈

**(55)** Assume $\neg s \,!\, s'$. So by definition (42):

(A) $\neg \exists p(s \models p \;\&\; s' \models \overline{p})$

We want to show $s' \oplus s^* = s^*$. By (9) and the fact that *HypeStates* encode only *Hype*-propositions (33), it suffices to show that $\forall p((s' \oplus s^*) \models p \equiv s^* \models p)$. So, by GEN, we show $(s' \oplus s^*) \models p \equiv s^* \models p$.

($\rightarrow$) Assume $(s' \oplus s^*) \models p$. Independently, by (38), we know:

$$\forall p((s' \oplus s^*) \models p \equiv s' \models p \vee s^* \models p)$$

Hence, $s' \models p \vee s^* \models p$. Assume, for reductio, that $\neg s^* \models p$. Then $s' \models p$, and since *Hype*-propositions are identical to their double negations (30), we know $s' \models \overline{\overline{p}}$. But it also follows from our reductio assumption, by (49), that $s \models \overline{p}$. So we've established $s \models \overline{p}\ \&\ s' \models \overline{\overline{p}}$. Existentially generalizing on $\overline{p}$, it follows that $\exists q(s \models q\ \&\ s' \models \overline{q})$, which contradicts (A).

(←) Exercise. ⋈

**(56)** Assume $s \trianglelefteq s'$. Since theorem (40) holds for any situations, it holds for *HypeStates*. So it follows that:

(A) $s \oplus s' = s'$

Now independently, by (54), we know that $s'$ is not incompatible with its Routley star image $s'^*$, i.e., $\neg s' \,!\, s'^*$. From this and (A), it follows that the fusion of $s$ and $s'$ is not incompatible with with the Routley star image of $s'$, i.e., that:

(B) $\neg(s \oplus s') \,!\, s'^*$

Now consider the following Lemma, which holds for any situations $s$, $s'$, and $s''$:

*Lemma*: $\neg(s \oplus s') \,!\, s'' \rightarrow \neg s \,!\, s''$

*Proof*: Assume $\neg(s \oplus s') \,!\, s''$. Then by definition of ! (42), we know $\neg\exists p((s \oplus s') \models p\ \&\ s'' \models \overline{p})$. Now suppose, for reductio, that $s \,!\, s''$. Then $\exists p(s \models p\ \&\ s'' \models \overline{p})$. Suppose $q_1$ is such a proposition, so that we know both $s \models q_1$ and $s'' \models \overline{q_1}$. But the former implies $s \oplus s' \models q_1$, by definition of $s \oplus s'$ (37). So we know $(s \oplus s') \models q_1\ \&\ s'' \models \overline{q_1}$. Hence, $\exists p((s \oplus s') \models p\ \&\ s'' \models \overline{p})$. Contradiction.

Given this Lemma, it follows from (B) that $s$ is not incompatible with $s'^*$, i.e., $\neg s \,!\, s'^*$. But by (55), this last result implies $s'^* \oplus s^* = s^*$. Hence, by (40), $s'^* \trianglelefteq s^*$. ⋈

## Bibliography

Barcan, R. C., 1946, "A Functional Calculus of First Order Based on Strict Implication", *The Journal of Symbolic Logic*, 11(1): 1–16.



Barwise, J., 1989, "Notes on Branch Points in Situation Theory", in J. Barwise, *The Situation in Logic*, CSLI Lecture Notes, No. 17, Stanford: Center for the Study of Language and Information Publications.

Barwise, J., and J. Perry, 1983, *Situations and Attitudes*, Cambridge, MA: MIT Press.

Berto, F., 2015, "A Modality called 'Negation'", *Mind*, 124(495): 761–793.

Berto, F., and G. Restall, 2019, "Negation on the Australian Plan", *Journal of Philosophical Logic*, 48: 1119–1144.

Copeland, B.J., 1979, "On When a Semantics Is Not a Semantics: Some reasons for disliking the Routley-Meyer semantics for relevance logic", *Journal of Philosophical Logic*, 8: 399–413.

Dunn, J. Michael, 1993, "Star and perp: Two treatments of negation", in James E. Tomberlin (ed.) *Philosophical Perspectives* (Volume 7), Atascadero, CA: Ridgeview Publishing Company, 331–357.

Enderton, H., 2001, *A Mathematical Introduction to Logic*, 2nd edition, San Diego: Harcourt/Academic Press.

Hintikka, J., 1959, "Towards a Theory of Definite Descriptions", *Analysis*, 19(4): 79–85.

Leitgeb, H., 2019, "HYPE: A System of Hyperintensional Logic (with an Application to Semantic Paradoxes)", *Journal of Philosophical Logic*, 48: 305–405.

Mares, E., 2004, *Relevant Logic: A Philosophical Interpretation*, Cambridge: Cambridge University Press.

Nolan, D., 1997, "Impossible Worlds: A Modest Approach", *Notre Dame Journal of Formal Logic*, 38(4): 535–572.

Odintsov, S., and H. Wansing, 2021, "Routley Star and Hyperintensionality", *Journal of Philosophical Logic*, 50: 33–56.

Punčochář, V., and I. Sedlár, 2022, "Routley Star in Information-Based Semantics", in A. Indrezejczak and M. Zawidzki (eds.), *Proceedings* *of the 10th International Conference on Non-Classical Logics: Theory and Applications* (NCL 2022), Electronic Proceedings in Theoretical Computer Science: 358, 285–397. doi:10.4204/EPTCS.358.21

Restall, G., 1995 [1999], "Negation in Relevant Logics (How I stopped worrying and learned to love the Routley Star)", Technical Report TR-ARP-3-95, Automated Reasoning Project, Research School of Information Sciences and Engineering, Australian National University; reprinted in *What is Negation?*, D. Gabbay and H. Wansing (eds.), Dordrecht: Kluwer, 1999, pp. 53–76. [Page reference is to the reprint.]

Restall, G., 2000, "Defining Double Negation Elimination", *Logic Journal of the IGPL*, 8(6): 853–860.

Routley, R., and R.K. Meyer, 1972, "The Semantics of Entailement III", *Journal of Philosophical Logic*, 1: 192–208.

———, 1973, "The Semantics of Entailment", in H. Leblanc (ed.), *Truth, Syntax, and Modality*, Amsterdam: North-Holland, 199–243.

Routley, R., and V. Routley, 1972, "The Semantics of First Degree Entailment", *Noûs*, 6(4): 335–369.

van Benthem, J.F.A.K., 1979, "What Is Dialectical Logic?", *Erkenntnis*, 14(3): 333–347.

Zalta, E., 1988, "Logical and Analytic Truths That Are Not Necessary", *The Journal of Philosophy*, 85(2): 57–74.

———, 1991, "A Theory of Situations", in J. Barwise, J. Gawron, G. Plotkin, and S. Tutiya (eds.), *Situation Theory and Its Applications*, Stanford: Center for the Study of Language and Information Publications, pp. 81–111.

———, 1993, "Twenty-Five Basic Theorems in Situation and World Theory", *Journal of Philosophical Logic*, 22: 385–428.

———, 1997, "A Classically-Based Theory of Impossible Worlds", *Notre Dame Journal of Formal Logic*, 38(4): 640–660.

———, m.s., *Principia Logico-Metaphysica*, URL = https://mally.stanford.edu/principia.pdf