An Axiomatic Theory of Truth

Gary M. Hardegree
Department of Philosophy
University of Massachusetts
Amherst, MA 01003
gmh@philos.umass.edu
http://www-unix.oit.umass.edu/~gmhwww

PART 1 – FORMAL DEVELOPMENT OF THEORY

1. Introduction

We propose to construct a formal theory[1] of truth, $\mathcal{T}$ (for 'truth'), which deals with the Liar Paradox in pretty much the same way that Axiomatic Set Theory deals with the Russell Paradox.

Just as Axiomatic Set Theory abandons the comprehension-schema of Naive Set Theory, and replaces it with a number of axioms that "do the same useful work", Axiomatic Truth Theory abandons the truth-schema of Naive Truth Theory, and replaces it with a number of axioms that "do the same useful work".

Obvious questions: What axioms replace the naive truth-schema? Can we abandon the naive truth-schema without throwing away any of its obviously correct and useful instances? Can we account for all the troublesome examples, including the liar, the truth-teller, contingently paradoxical sentences, etc.

These questions can only be answered by actually developing the theory.

2. Syntactic Apparatus for $\mathcal{T}$

The basic logical apparatus of $\mathcal{T}$ is multi-sorted free first-order logic (without descriptions) augmented by two modes of sentential quotation – *direct-quotation* and *indirect-quotation*.

(1)	quote…unquote	[direct-quotation]
(2)	that …	[indirect-quotation]

Grammatically, these are both 1-place subnectives, which means that each one takes a single sentence (S) and delivers a noun phrase (N), in particular a singular term. Superficially, direct-quotation and indirect-quotation are very similar. At a deeper level, however, they are importantly different. Specifically, whereas indirect-quotation is *completely transparent*, direct-quotation is *completely opaque*.[2] This has important consequences for the formal development of $\mathcal{T}$.

Note that, $\mathcal{T}$ posits only sentential quotation. It is left for future research to develop logical extensions of $\mathcal{T}$ that posit direct-quotation (*and* indirect-quotation) of other grammatical expressions.

---

[1] Although I refer to $\mathcal{T}$ as a theory of truth, I propose it as a *logic* of truth. See Section (24). *Perhaps* it is "only" a *theory of truth*. Basically, at least as a matter of usage, until/unless a formal system pertaining to a linguistic domain (e.g., connectives, quantifiers, identity, descriptions, etc.) is pretty much universally accepted, it is a theory rather than a logic. In this regard, consider Frege's theory of descriptions, Russell's theory of descriptions, world theory, and counterpart theory.

[2] Direct-quotation is *categorially opaque*. Categorial opacity means that a functor cannot be combined with other functors to produce derivative functors. Upon being inserted into direct-quotation, an expression, no matter how grammatically complex, is "squashed down" to its fundamental literal material; no grammatical items, not even words, retain their separate identities. Categorial opacity implies semantic (intensional) opacity which in turn implies referential (denotational) opacity. Opacity pertains to inter-substitution of equivalent items. Semantic opacity pertains to inter-substitution of items with the same intension (sense), referential opacity pertains to substitution of items with the same extension (denotation). Although Bush is (identical to) the U.S. president in 1990, the word 'Bush' is not identical to the noun phrase 'the U.S. president in 1990'. Also, although 'bachelor' *means* 'eligible but unmarried male', in which case these two expressions have the same intension, the word 'bachelor' is obviously not identical to the phrase 'eligible but unmarried male'.

In addition to direct and indirect quotation, $\mathcal{T}$ also posits a number of *logical* expressions with ordinary categories. In the following, we present the complete list, including repeating items (1) and (2).

	Expression	**Category**	**Symbol**
(1)	quote…unquote	1-place subnective	$\ulcorner\ldots\urcorner$
(2)	that …	1-place subnective	$\langle\ldots\rangle$
(3)	… is a sentence	1-place predicate	$\mathbb{S}$
(4)	… is a proposition	1-place predicate	$\mathbb{P}$
(5)	… is true	1-place predicate	$\mathbb{T}$
(6)	… is normal	1-place predicate	$\mathbb{N}$
(7)	… implies …	2-place predicate[3]	$\precsim$
(8)	the negation of …	1-place function sign[4]	$\neg$
(9)	the conjunction of … and …	2-place function sign[5]	$\wedge$

In addition to these primitive expressions (concepts), $\mathcal{T}$ also posits a number of defined expressions (concepts), including the following.

(10)	… is an ordinary object	1-place predicate	$\mathbb{O}$
(11)	… is false	1-place predicate	$\mathbb{F}$
(12)	… is gapped	1-place predicate	$\mathbb{G}$
(13)	… says …	2-place predicate	[none]
(14)	… is equivalent to …	2-place predicate[6]	$\approx$
(15)	… entails …	2-place predicate[7]	$\leqslant$
(16)	… is strictly-equivalent to …	2-place predicate	$\equiv$
(17)	… says … and nothing more	2-place predicate	[none]
(18)	the necessary proposition	proper noun	1
(19)	the impossible proposition	proper noun	0
(20)	strict-necessity operator	1-place connective	$\Box$
(21)	weak-necessity operator	1-place connective	$\boxdot$
(22)	entailment operator	2-place connective	$\leqslant$ [8]

[3] When I say that 'implies' is a predicate, I mean it. In particular, it takes two noun phrases (N2) and delivers a sentence (S). The implication-predicate 'implies' should not to be confused (in the unfortunate, but venerable, tradition of Russell) with the 'if…then…' connective, which is a two-place *connective*, which means of course that it takes two sentences (S2) and delivers a sentence (S).

[4] Once again, we *must* be grammatically fastidious. The expression 'the negation of …' is a one-place *function sign*, which means that it takes a noun phrase (N) and delivers a noun phrase (N). This should not be confused with the one-place connective 'not…', which being a one-place connective takes a sentence (S) and delivers a sentence (S).

[5] See previous footnote about function signs.

[6] See footnote 3. Note that this is a two-place predicate, not a two-place connective.

[7] This is a stronger version of the ordinary 'implies'. Note that it is a predicate, not a connective.

(23)	normalcy operator	1-place connective	$\mathcal{N}$
(24)	strict-equivalence operator	2-place connective	$=$ [9]

3. Sortal Assumptions

Theory $\mathcal{T}$ posits three basic categories (*sorts*) of objects.

(1) sentences
(2) propositions
(3) "ordinary" objects

It is moreover presumed that these are mutually exclusive and jointly exhaustive. In other words, the following are *presumed* to be *logically* true.

(1) every *thing* is a sentence, a proposition, or an ordinary object;
(2) no sentence is a proposition;
(3) no sentence or proposition is an ordinary object.

Given that the intended domain of $\mathcal{T}$ neatly divides into three sorts, it is convenient to use sortal variables, which we present in the following table.

x, y, z, x_1, y_1, etc.	things in general
s, s_1, s_2, etc.	sentences
p, q, r, p_1, q_1, etc.	propositions
o_1, o_2, etc.	ordinary objects

[8] In order to avoid symbolic over-kill, we frequently use the same symbol for many different uses. For example, '$\leqslant$' is both a predicate and a connective, as well as a relation in the semantics (Part 2). Hopefully, context will determine which use is intended. If we are not careful, however, we will make use-mention errors. If we write '$\phi \leqslant \psi$', we might be lazy and read it as "ϕ entails ψ", but we really mean "that(ϕ) entails that(ψ)". See Section 9.
[9] Using '=' as a connective – in addition to its primitive logical use – can cause confusion, so we must be careful. We write, for example, '$\phi=\psi$' but we should read it "that(ϕ) = that(ψ)". The logic of the '=' connective is herein called Sentential Identity Logic (SIL). Its principles are fairly obvious, and are nearly identical in appearance to ordinary identity logic(IL). The only difference concerns substituting "identical" formulas within variable-binding operators. In particular, we cannot infer '$\forall xFx = \forall xGx$' from '$Fx = Gx$'.

4. Non-Logical Vocabulary; the Family of $\mathcal{T}$-Languages

The formal system $\mathcal{T}$ is *proposed* as a logic in exactly the same way that first-order logic is proposed as a logic. In particular, associated with each is a corresponding class of languages and theories. In the case of first-order logic, these languages/theories are called (not surprisingly) first-order languages/theories. In the case of $\mathcal{T}$, we propose to call the associated family of languages $\mathcal{T}$-languages/theories.

The definition of $\mathcal{T}$-language is very similar to the definition of a first-order language. In particular, a $\mathcal{T}$-language is a formal language specified by a vocabulary and rules of formation, given as follows.

1. Vocabulary
   - 1.1. logical vocabulary (as given above)
   - 1.2. non-logical vocabulary
     - 1.2.1. zero or more proper nouns
     - 1.2.2. zero or more n-place predicates (for n = 0, 1, 2, …)[10]
     - 1.2.3. zero or more n-place function signs (for n = 1, 2, …)[11]

Note carefully that although $\mathcal{T}$ posits two *logical* subnectives – direct-quotation and indirect-quotation – it does not admit *non-logical* subnectives.[12]

2. Rules of Formation
   Since every expression is assigned a category, and since we identity category S with formulas, and category N with singular terms, the rules of formation for a $\mathcal{T}$-language are completely determined by the principles of categorial grammar. There is no need to write down the rules of formation explicitly.[13]
   However, we do write down a few examples, as follows.[14]
   - 2.1. if ϕ_1 is a formula, and ϕ_2 is a formula, then $(\phi_1\&\phi_2)$ is a formula;
   - 2.2. if τ_1 and τ_2 are singular terms, then $(\tau_1 = \tau_2)$ is a formula;
   - 2.3. if ϕ is a formula, then $\ulcorner\phi\urcorner$ is a singular term;[15]

---

[10] Note that we admit zero-place predicates, so that the sentence letters ('P', 'Q' , etc.) of sentential logic count as admissible sentences of $\mathcal{T}$. Such sentences correspond to subject-less sentences like 'it is raining'. To see that 'it is raining' is subject-less, ask yourself how you would do its semantics in first-order logic treating 'it' as a genuine noun phrase.

[11] Although we admit zero-place predicates, we do not admit zero-place function signs *per se*. What we have instead are *proper nouns*, which perform the same grammatical duty as zero-place function signs. Of course, we could banish proper nouns instead, but it seems more terminologically natural to posit proper nouns (e.g., '0' in Peano Arithmetic) and banish zero-place function signs.

[12] Nor does it admit *non-logical* connectives (except *implicitly* – e.g., attitude and modal "connectives"), nor does it admit non-logical quantifiers (except *implicitly* – e.g, 'every F is such that …'). [[Is there a genuine difference between explicit and implicit connectives? Yes! The proposed semantics for $\mathcal{T}$ (see below) has *no* means of interpreting *atomic* non-logical connectives and quantifiers, although it obviously is capable of interpreting implicit connectives and quantifiers.]] Finally, $\mathcal{T}$ does not admit higher-order expressions (for example, plurals, adverbs, pro-sentential variables, pro-sentential quantifiers). Expanding the theory in these ways is left for future research.

[13] Quantifiers are categorially special: each one takes an individual variable (V) plus a formula (S) and delivers a formula (S).

[14] Notice that parentheses are officially part of the rules of formation. In particular, every two-place infix functor (e.g., &, =, says, implies) comes with parentheses in its formation rule. Of course, in actual practice, outer parentheses are frequently dropped when they do not contribute to parsing a formula. For example, compare '∀x(x=x)' and 'x=x'.

[15] See Section 5 for important information about corner-quotes. We do not use corner-quotes in the Quine-specified manner.

2.4. if ϕ is a formula, then $\langle\phi\rangle$ is a singular term;

2.5. if τ_1 and τ_2 are singular terms, then $(\tau_1 \precsim \tau_2)$ is a formula;

etc.

Note carefully, there is no restriction on exactly what the non-logical items of a $\mathcal{T}$-language are, so long as they are categorially admissible. We have in mind a logic for a fragment of ordinary language. Therefore, for the purposes of discussion, we presume that the formal language we are discussing, which we call $\mathcal{L}$, contains every first-order expression of ordinary English (*cum* category-tag). If an expression of ordinary English is categorially ambiguous, as many are, it will give rise to multiple instantiations in the corresponding formal language $\mathcal{L}$. In particular, every ordinary language expression must be "tagged" with an explicit unique grammatical category, and this category must be admissible. Of course, in the latter regard, the $\mathcal{T}$-framework is insistently and fastidiously (and perhaps stubbornly and annoyingly) *first-order*. Non-logical expressions *must* be proper nouns, functions signs, or predicates.[16]

5. Meta-Linguistic Devices; an Important Note on Corner-Quotes

Our presentation of the theory employs various meta-linguistic devices, examples of which you have already seen. For example, Greek letters are used as *schematic variables* ranging over various items in the given $\mathcal{T}$-language $\mathcal{L}$. Schematic variables are, of course, critical for presenting the rules of inference and axiom schemata.

You will no doubt have noticed that $\mathcal{T}$ employs corner-quotes, which we now explain.

As prescribed by Quine, corner-quotes are employed exclusively in the meta-language as a way of collecting a series of meta-linguistic noun phrases and forming a compound noun phrase. The following example illustrates how Quine-quotes can be used.

if ε_1 is a singular term, and ε_2 is a singular term, then $\ulcorner \varepsilon_1 = \varepsilon_2 \urcorner$ is a formula.

You will readily see that dropping the corner quotes does not render the sentence unintelligible, but it does make it somewhat harder to parse visually.

We do *not* use corner quotes in the manner prescribed by Quine [i.e., as a method of constructing compound noun phrases in the meta-language].

[16] Notice, in this connection, that this eliminates expressions like 'and' 'or' and 'not' – *qua* connectives. This is ok, since our language already includes (counterparts of) these in its logical vocabulary. There can be no atomic non-logical connectives in a $\mathcal{T}$-language or $\mathcal{T}$-theory. On the other hand, derivative (molecular) connectives are admissible. For example, if the language $\mathcal{L}$ has the verb 'believes' and the proper noun 'Max', then there will automatically be a derivative one-place connective 'Max believes that …' Derivative connectives are not *purely* logical, although they may be quasi-logical. A *theory* of belief may stipulate, in effect, that 'Max believes that' is a modal operator satisfying the principles of modal system K (or whatever). Also there is no reason that there cannot be *theories of truth* under-written by logical system $\mathcal{T}$. For example, a *theory* of truth might posit that nothing is true except insofar as God wills it so. However, this posit would not be a logical truth.

Our technique of producing compound noun phrases in the meta-language is to use a combination of *ambiguous symbols* and *direct-concatenation*.[17]

First, we assume that every object-language symbol, regardless of its category in the object language, also appears as a symbol in the meta-language, the key difference being that the symbol in the meta-language is a proper noun, which denotes that object language symbol.

Second, if a series $\eta_1,\ldots,\eta_k$ of noun phrases occurs in the meta-language, and denotes a corresponding series $\varepsilon_1,\ldots,\varepsilon_k$ of object-language expressions, then the result of concatenating $\eta_1,\ldots,\eta_k$ is, by definition, a compound noun phrase of the meta-language that denotes the corresponding concatenation of the object language expressions $\varepsilon_1,\ldots,\varepsilon_k$.

Corner-quotes are *used* ambiguously, both in the object language as direct-quotation marks, **and** in the metalanguage as *proper nouns* denoting those quote-marks.

The advantage of this notational choice is that it allows us to use ordinary quote-marks in their customary manner – both formally, and colloquially – to name the literal material within them. The following illustrates this.

The suggested reading of '$\ulcorner$' is "quote", and the suggested reading of '$\urcorner$' is "un-quote". On the other hand, for the sake of brevity, the suggested reading of '$\ulcorner\phi\urcorner$' is "quote-phi", which is analogous to reading '$\langle\phi\rangle$' as "that-phi"

[17] If we were to use quote-plus notation (as in the programming language Basic) instead of direct-concatenation, then we would never have problems understanding the true grammar of meta-linguistic expressions, but unfortunately they would be exceptionally clumsy. The following compare a number of sentences employing the two notational schemes.

(1) if ϕ is a formula, then ϕ is a formula;
(1′) if ϕ is a formula, then ' ' + ϕ is a formula;
(2) if ε_1 and ε_2 are singular terms, then ε_1=ε_2 is a formula;
(2′) if ε_1 and ε_2 are singular terms, then ε_1 + '=' + ε_2 is a formula;
(3) if ϕ is a formula, then $\ulcorner\phi\urcorner$ is a singular term;
(3′) if ϕ is a formula, then '$\ulcorner$' + ϕ + '$\urcorner$' is a singular term.

6. Initial Axioms of $\mathcal{T}$

We now turn to the formal development of the content of $\mathcal{T}$. Our first four items make formally explicit our sortal assumptions about sentences, propositions, and ordinary objects.

(a1) ***x*** is a sentence $\rightarrow$ ∃y[***x*** = y]

(a2) ***x*** is a proposition $\rightarrow$ ∃y[***x*** = y]

(a3) ~∃x{x is a sentence & x is a proposition}

(d1) ***x*** is an ordinary object $=_{df}$ ∃y{***x***=y & y is a not a sentence & y is not a proposition}

From these, we can immediately obtain the following simple theorem.

(t1) x is a sentence **XOR** x is a proposition **XOR** x is an ordinary object[18]

The next axiom tells us what sorts of objects *can* be true.

(a4) ***x*** is true $\rightarrow$. ***x*** is a sentence **OR** ***x*** is a proposition

In other words,

The only things that can be true are sentences and propositions.

The next two principles are obvious, but must nonetheless be officially postulated.

(a5) $\ulcorner\phi\urcorner$ is a sentence

(a6) $\langle\phi\rangle$ is a proposition

From these we get the following obvious theorems.

(t2) $\exists x\{x = \ulcorner\phi\urcorner\}$

(t3) $\exists x\{x = \langle\phi\rangle\}$

(t4) $\langle\phi\rangle \neq \ulcorner\phi\urcorner$

---

[18] Here, 'xor' is exclusive 'or'. Note carefully that, despite its superficial appearance here, 'xor' is an *anadic* connective. This is because dyadic-xor does not correctly yield triadic-xor, tetradic-xor, etc.

7. Negation, Conjunction; Boolean Algebras

In the present section, we discuss some of the simplest axioms about sentences and propositions. One thing we wish to achieve in the formal development of $\mathcal{T}$ is to show that propositions form (at least) a Boolean algebra. In order to formulate this notion in the object language of $\mathcal{T}$, we introduce two additional primitive function signs – *negation* and *conjunction*.[19]

7.1. Negation

The theory posits a primitive one-place *function sign* for negation – symbolized by '¬' – whose characteristic axioms are given as follows. The first axiom tells us exactly what sorts of objects have negations.

(a7) $\exists y[y = \neg \boldsymbol{x}]$ $\leftrightarrow$. $\boldsymbol{x}$ is a sentence **OR** $\boldsymbol{x}$ is a proposition

In other words, sentences have negations, propositions have negations, but ordinary objects do not have negations.

The next axiom schema tells us that the negation of $\ulcorner\phi\urcorner$ is exactly what you would expect it to be.

(a8) $\neg\ulcorner\phi\urcorner = \ulcorner{\sim}\phi\urcorner$

Replacing direct-quote marks by indirect-quote marks yields the following corresponding propositional axiom schema.

(a9) $\neg\langle\phi\rangle = \langle{\sim}\phi\rangle$

7.2. Conjunction

The theory also posits a primitive *function sign* for conjunction – symbolized by '∧' – whose characteristic axioms are given as follows.

(a10) $\exists z[z = \boldsymbol{x}\wedge\boldsymbol{y}]$ $\leftrightarrow$: $\mathbb{S}\boldsymbol{x}$ & $\mathbb{S}\boldsymbol{y}$.**OR**. $\mathbb{P}\boldsymbol{x}$ & $\mathbb{P}\boldsymbol{y}$

(a11) $\ulcorner\phi_1\urcorner \wedge \ulcorner\phi_2\urcorner = \ulcorner\phi_1 \,\&\, \phi_2\urcorner$

(a12) $\langle\phi_1\rangle \wedge \langle\phi_2\rangle = \langle\phi_1 \,\&\, \phi_2\rangle$

7.3. The Remaining Boolean Operations

We can define the remaining Boolean operations in the expected manner.

(d2) $\boldsymbol{p} \vee \boldsymbol{q} =_{\text{df}} \neg(\neg\boldsymbol{p} \wedge \neg\boldsymbol{q})$

(d3) $\boldsymbol{p} \rightarrow \boldsymbol{q} =_{\text{df}} \neg(\boldsymbol{p}_1 \wedge \neg\boldsymbol{q})$

(d4) $\boldsymbol{p} \leftrightarrow \boldsymbol{q} =_{\text{df}} (\boldsymbol{p} \wedge \boldsymbol{q}) \vee (\neg\boldsymbol{p} \wedge \neg\boldsymbol{q})$

[19] As usual, we have to pay close attention to grammatical categories here. These are function signs, not connectives. Grammatically, they take noun phrases and deliver noun phrases.

7.4. Boolean Algebras

The next axiom group ensures that the propositions form a Boolean algebra.

(a13) A set of equations that completely (and consistently) describe Boolean algebras; any set will do. [An example is given in Part 3.]

8. Entailment

8.1. Entailment Predicate for Propositions

Every Boolean algebra comes equipped with an *order relation*, $\leqslant$, defined as follows.

$$x \leqslant y \quad =_{df} \quad x = x \wedge y$$

We accordingly introduce '$\leqslant$' as a defined *predicate*, which we call *entailment*.

(d5) $\boldsymbol{p} \leqslant \boldsymbol{q} \quad =_{df} \quad \boldsymbol{p} = \boldsymbol{p} \wedge \boldsymbol{q}$

It is routine to show that $\leqslant$ is a *partial-ordering* of the propositions, which means that it satisfies the following.

(t5) $p \leqslant p$ [$\leqslant$ is reflexive]

(t6) $p \leqslant q \;\&\; p \leqslant r \;.\!\rightarrow\; p \leqslant r$ [$\leqslant$ is transitive]

(t7) $p \leqslant q \;\&\; q \leqslant p \;.\!\rightarrow\; p = q$ [$\leqslant$ is anti-symmetric]

It also satisfies the following principles.

(t8) $p \leqslant q \wedge r \leftrightarrow\!.\; p \leqslant q \;\&\; p \leqslant r$ [*glb* property]

(t9) $p \leqslant q \rightarrow \neg q \leqslant \neg p$ [contraposition]

8.2. Entailment For Sentences

Many propositional concepts can be naturally extended to sentences, and entailment is no exception. What can be confusing, however, is that we can define both an entailment *predicate*, and a entailment *connective*. In order to clarify the grammatical and conceptual issues, we extensively discuss these ideas in the next section.

9. Entailment and Necessity

First, every Boolean algebra has (admits) two privileged elements – 0 and 1 – which may be defined as follows.

(d6) $0 \quad =_{df} \quad \boldsymbol{p} \wedge \neg \boldsymbol{p}$

(d7) $1 \quad =_{df} \quad \neg 0$

The well-definedness of 0, and hence 1, depends upon the following supporting theorem.

(t10) $\forall p \forall q\{p\wedge\neg p = q\wedge\neg q\}$

We will call 1 the *necessary proposition*, and we will call 0 the *impossible proposition*.[20] We also add the following fairly obvious principles – the necessary proposition is true, and impossible proposition is not true.

(a14) $\mathbb{T}[1]$

(a15) $\sim\mathbb{T}[0]$

9.1. Moving From First to Second Grade Modal Involvement

In the current section, we show how to move from Quine's first grade of modal involvement to Quine's second grade of modal involvement. We first present the technique informally, after which we offer the formal definitions.

Given the necessary proposition 1, we next *informally* define a corresponding necessity *predicate* as follows.

(*id*1) $\mathbb{N}[\boldsymbol{p}] \quad =_{df} \quad \boldsymbol{p} = 1$

Here '$\mathbb{N}[\boldsymbol{p}]$' is read "***p*** is necessary". This allows us to define *informally* a necessity *connective* as follows.

(*id*2) $\Box\phi \quad =_{df} \quad \mathbb{N}\langle\phi\rangle$

In other words,

> $\Box\phi$
>
> is shorthand for:
>
> that(ϕ) is necessary

By a simple grammatical transformation, the latter expression is equivalent to:

> "it" is necessary that ϕ

Here, the "it" is the "it" found so puzzling by the duck in *Alice in Wonderland* ("The Caucus Race"). In our particular case, the pronoun 'it' takes the noun phrase 'that ϕ' as its pronominal antecedent. Accordingly,

> $\Box\phi$ means: it is necessary that ϕ.

[20] In the special case of simple two-valued logic, there are only two propositions – the "true" (T) and the "false" (F). In this case, there are no modal distinctions – the necessary proposition is just T, and the impossible proposition is just F.

The previous analysis was informal. The formal analysis preserves the spirit of the informal analysis; it simply dispenses with the necessity predicate entirely, and simply defines the box-operator directly, as follows.

(d8) $\Box\phi =_{df} \langle\phi\rangle = 1$

9.2. The Truth Operator (Connective)

Next, we note that, just as the necessity predicate can be re-rendered as a necessity operator, the truth *predicate* can be re-rendered as a truth *operator* (connective), as follows.

(d9) $\mathbb{T}\phi =_{df} \quad \mathbb{T}\langle\phi\rangle$

Note carefully that the same symbol is used for the primitive truth-predicate and the derivative truth-operator. Which one is meant in a given circumstance will hopefully be apparent from context.

In other words,

$\mathbb{T}\phi$ is read: that(ϕ) is true

which is a transformational variant of:

it is true that ϕ

9.3. Entailment Operator

Just as with necessity and truth, entailment can also be rendered as a connective. In particular, we can define an entailment *operator* (i.e., connective) as follows, which is followed by an immediate theorem.

(d10) $\phi_1 \leqslant \phi_2 \quad =_{df} \quad \langle\phi_1\rangle \leqslant \langle\phi_2\rangle$

(t11) $\phi_1 \leqslant \phi_2 \;\leftrightarrow\; \Box(\phi_1 \rightarrow \phi_2)$[21]

Note carefully: in order to reduce the number of symbols,
we employ '$\leqslant$' for both the *connective* and the *predicate*.
Context will determine which one is intended.

The basic idea is that when it *appears* that we are committing a sophomoric category error, or confusing *use* and *mention*, actually we are simply trying to minimize symbolic clutter. In particular, when it appears that a predicate has been applied to sentences, it has really been applied to the nominalized counterparts of those sentences. In particular,

P "entails" Q

which is admittedly ungrammatical, *really is shorthand for:*

[21] This is, course, exactly how C.I. Lewis proposed that the strict-implication connective be related to the necessity connective.

that(P) entails that(Q)[22]

9.4. Entailment-Predicate For Sentences

The entailment-connective is one way to transfer the order-relation $\leqslant$ to sentences. The other way is simply to construct a corresponding order-relation (predicate) on sentences. It is officially defined as follows.

(d11) $\boldsymbol{s}_1 \leqslant \boldsymbol{s}_2 \quad =_{\text{df}} \quad \langle \mathbb{T}\boldsymbol{s}_1 \rangle \leqslant \langle \mathbb{T}\boldsymbol{s}_2 \rangle$

The following is an immediate theorem.

(t12) $\text{s}_1 \leqslant \text{s}_2 \quad \leftrightarrow \quad \mathbb{T}\text{s}_1 \leqslant \mathbb{T}\text{s}_2$

In other words,

Sentence $\boldsymbol{s}_1$ entails sentence $\boldsymbol{s}_2$ iff the proposition that($\boldsymbol{s}_1$ is true⟩ entails the proposition that($\boldsymbol{s}_2$ is true)

9.5. Entailments between Sentences and Propositions

So far, the entailment *relation* holds among sentences, and among propositions, but not among sentences-*and*-propositions. This is remedied by adding the following further principles.

(d12) $\boldsymbol{s} \leqslant \boldsymbol{p} \quad =_{\text{df}} \quad \langle \mathbb{T}\boldsymbol{s} \rangle \leqslant \boldsymbol{p}$

(d13) $\boldsymbol{p} \leqslant \boldsymbol{s} \quad =_{\text{df}} \quad \boldsymbol{p} \leqslant \langle \mathbb{T}\boldsymbol{s} \rangle$

Note, however, that we do not allow any mixed-category expressions like the following.

✕ $\quad \text{x} \leqslant \phi; \phi \leqslant \text{x};$

Specifically, we use '⩽' as a predicate (primitive) and as a connective (derivative), but that is it!

9.6. Strict Equivalence

Once we have the (trans-sortal) entailment *predicate* ⩽. we can define an associated strict-equivalence *predicate*, ≡, as follows.

(d14) $\boldsymbol{x} \equiv \boldsymbol{y} \quad =_{\text{df}} \quad \boldsymbol{x} \leqslant \boldsymbol{y} \ \& \ \boldsymbol{y} \leqslant \boldsymbol{x}$

---

[22] In this connection, see "Grammatical Propaedeutic" in Anderson and Belnap's *Entailment: The Logic of Relevance and Necessity*. Of course, the trouble with our little bit of grammatical therapy is that it renders nested implication statements nearly impossible to understand. For example,

$\phi_1 \leqslant (\phi_2 \leqslant \phi_3)$ means: that(ϕ_1) entails that(that(ϕ_2) entails that(ϕ_3))

$(\phi_1 \leqslant \phi_2) \leqslant \phi_3$ means: that(that(ϕ_1) entails that(ϕ_2)) entails that(ϕ_3)

It is easy to show that the strict-equivalence relation is indeed an equivalence relation over the appropriate class of objects (i.e., sentences and propositions).

(t13) $s \equiv s \,;\, p \equiv p$

(t14) $x \equiv y \;\rightarrow\; y \equiv x$

(t15) $x \equiv y \;\&\; y \equiv z \;.\!\rightarrow\; x \equiv z$

We note that, when restricted to propositions, strict-equivalence is just identity.

(t16) $p \equiv q \;\leftrightarrow\; p = q$

Of course, we do not have a corresponding theorem for sentences.

Finally, we can define a strict-equivalence *connective* in the same manner that we defined the entailment connective.[23]

(d15) $\phi_1 = \phi_2 \quad =_{df} \quad \langle\phi_1\rangle = \langle\phi_2\rangle$

We can then prove the following theorems.

(t17) $\phi_1 = \phi_2 \;\leftrightarrow\; \langle\phi_1\rangle \equiv \langle\phi_2\rangle$

(t18) $\phi_1 = \phi_2 \;\leftrightarrow\; \Box\{\phi_1 \leftrightarrow \phi_2\}$

(t19) $s_1 \equiv s_2 \;\leftrightarrow\; \mathbb{T}s_1 = \mathbb{T}s_2$

9.7. Entailment and Quantification

Before we continue, we list a key axiom schema about how entailment is related to quantification (Quine's third grade of modal involvement).

(a16) $p \leqslant \langle\forall\nu\phi\rangle \;\leftrightarrow\; \forall\nu\{p \leqslant \langle\phi\rangle\}$

Here, ϕ is any formula, and ν is any variable distinct from 'p'. The following is a simple instance.

(t20) $p \leqslant \langle\forall xFx\rangle \;\leftrightarrow\; \forall x\{p \leqslant \langle Fx\rangle\}$

In other words,

p entails that everything is F
if and only if
for any *thing x*, p entails that *x* is F

We also have as a special case the following theorem schema – the well-known Barcan-formula.

(t21) $\Box\forall\nu\phi \;\leftrightarrow\; \forall\nu\Box\phi$

[23] Once again, note the potential for confusion. 'ϕ=ψ' does not mean "ϕ is ψ", but rather "that(ϕ) is that(ψ)".

10. Implication

Most logical systems have exactly one *order relation*, which is given by the partial-order relation $\leqslant$ on the algebra of semantic values.[24] By contrast, system $\mathcal{T}$ posits an additional *primitive relation* $\precsim$, which is called *implication*, and which is generally weaker than $\leqslant$. The role of $\precsim$ is to capture peculiar ("paradoxical") implications, such as those engendered by the liar sentence.

10.1. Simple Axioms

The first axiom gives us basic category information.

(a17) $\boldsymbol{x} \precsim \boldsymbol{y} \to.\ \mathbb{S}\boldsymbol{x}$ OR $\mathbb{P}\boldsymbol{x}$.&. $\mathbb{S}\boldsymbol{y}$ OR $\mathbb{P}\boldsymbol{y}$

In other words, implication holds only among sentences and propositions. Notice in particular that, like entailment, implication is trans-sortal — a sentence can imply a proposition, and a proposition can imply a sentence.

The next axiom tells that entailment is at least as strong as implication. Every instance of entailment is also an instance of implication.

(a18) $x \leqslant y \to x \precsim y$

Since, the entailment relation $\leqslant$ is reflexive over its domain, so is the implication relation $\precsim$. Specifically, we have the following theorem.

(t22) $s \precsim s$; $p \precsim p$

The next axiom tells us that $\precsim$ is transitive, so in connection with (t22), we have that $\precsim$ is a quasi-ordering relation over sentences and propositions.

(a19) $x \precsim y$ & $y \precsim z .\to x \precsim z$

The critical difference between implication and entailment is that implication is *not* postulated to be anti-symmetric for propositions. In other words, we do not have the following theorem.

✕ $p \precsim q$ & $q \precsim p .\to p = q$

Nevertheless, given that $\precsim$ is a quasi-ordering, we can define an affiliated equivalence-predicate as follows.

(d16) $\boldsymbol{x} \approx \boldsymbol{y} \quad =_{\mathrm{df}} \quad \boldsymbol{x} \precsim \boldsymbol{y}$ & $\boldsymbol{y} \precsim \boldsymbol{x}$

In other words, equivalence is just mutual implication. It is then easy to prove the following theorems, which collectively tell us that $\approx$ is an equivalence relation over the appropriate domain.

(t23) $s \approx s$; $p \approx p$

[24] For example, in traditional two-valued logic, the two "values" are T and F. The "implication" ordering is then "material implication" [i.e., F $\leqslant$ F $\leqslant$ T $\leqslant$ T]. Accordingly, when we say that P "materially implies" Q, we "really" mean that the truth-value of P implies [$\leqslant$] the truth value of Q.

(t24) $x \approx y \rightarrow y \approx x$

(t25) $x \approx y \ \& \ y \approx z \ .\!\rightarrow x \approx z$

10.2. Axioms to Ensure that ≈ is a Congruence Relation

As seen in the previous section, ≈ is an equivalence relation. We would like something stronger, that ≈ is a Boolean *congruence*. For this reason, we introduce the following further axioms.

(a20) $x \precsim y \rightarrow \neg y \precsim \neg x$

(a21) $x \precsim p_1 \ \& \ x \precsim p_2 \ .\!\rightarrow x \precsim p_1 \wedge p_2$

(a22) $x \precsim s_1 \ \& \ x \precsim s_2 \ .\!\rightarrow x \precsim s_1 \wedge s_2$

With these in hand, we can prove the following theorems.

(t26) $s_1 \precsim s_2 \leftrightarrow s_1 \approx s_1 \wedge s_2$

(t27) $p_1 \precsim p_2 \leftrightarrow p_1 \approx p_1 \wedge p_2$

(t28) $x \approx y \rightarrow \neg x \approx \neg y$

(t29) $s_1 \approx s_2 \ \& \ s_3 \approx s_4 \ .\!\rightarrow s_1 \wedge s_3 \approx s_2 \wedge s_4$

(t30) $p_1 \approx p_2 \ \& \ p_3 \approx p_4 \ .\!\rightarrow p_1 \wedge p_3 \approx p_2 \wedge p_4$

10.3. Second-Grade Modalities

Just as we can use the entailment *predicate* to define a strict-implication *connective*, we can do the same thing with implication. In particular, we have the following definitions, and theorem.

(d17) $\boxdot\phi =_{df} \langle\phi\rangle \approx 1$

(d18) $\phi_1 \precsim \phi_2 =_{df} \langle\phi_1\rangle \precsim \langle\phi_2\rangle$

(t31) $\phi_1 \precsim \phi_2 \ .\!\leftrightarrow \boxdot\{\phi_1 \rightarrow \phi_2\}$

Notice, once again, that we use a single symbol '≾' for both the connective and the predicate.

10.4. Implication and Quantification

Recall how entailment and quantification are related (Section 9.7). Implication and quantification are similarly related.

(a23) $p \precsim \langle\forall\nu\phi\rangle \leftrightarrow \forall\nu\{p \precsim \langle\phi\rangle\}$

Here, ϕ is any formula, and ν is any variable distinct from 'p'. The following are both theorems.

(t32) $p \precsim \langle \forall x Fx \rangle \leftrightarrow \forall x\{p \precsim \langle Fx \rangle\}$

(t33) $\boxdot \forall \nu \phi \leftrightarrow \forall \nu \boxdot \phi$

11. The Relation Between Entailment, Implication, and Truth

We next examine the relation between our two order relations ($\leqslant$ and $\precsim$) and truth. We begin with the following fundamental principles.

(a24) $x \leqslant y \rightarrow . \ \mathbb{T}x \rightarrow \mathbb{T}y$ [$\leqslant$-OUT]

(a25) $x \precsim y \rightarrow . \ \mathbb{T}x \rightarrow \mathbb{T}y$ [$\precsim$-OUT]

In other words, if x entails/implies y, and x is true, then y is true. No big surprise here.

These can be strengthened as follows.

(a26) $x \leqslant y \leftrightarrow \Box(\mathbb{T}x \rightarrow \mathbb{T}y)$

(a27) $x \precsim y \leftrightarrow \boxdot(\mathbb{T}x \rightarrow \mathbb{T}y)$

In addition to these two modal principles, $\mathcal{T}$ also postulates the following important axiom schemata, which is followed by an immediate corollary.

(a28) $\phi_1 \leqslant \phi_2 \rightarrow \mathbb{T}\ulcorner\phi_1\urcorner \leqslant \mathbb{T}\ulcorner\phi_2\urcorner$ [monotonicity]

(t34) $\phi_1 = \phi_2 \rightarrow \mathbb{T}\ulcorner\phi_1\urcorner = \mathbb{T}\ulcorner\phi_2\urcorner$ [SIL]

Note carefully that, although the entailment relation $\leqslant$ is monotonic, the (weaker) implication relation $\precsim$ is not.

✕ $\phi_1 \precsim \phi_2 \rightarrow \mathbb{T}\ulcorner\phi_1\urcorner \precsim \mathbb{T}\ulcorner\phi_2\urcorner$

12. Modal Principles

We now take this opportunity to employ the modal operators to state some other important principles of $\mathcal{T}$.

12.1. Rules of Inference

We begin with some important rules of inference.[25]

(r1) $\{\vdash\phi\} \hookrightarrow \boxdot\phi$

[25] The rules in question are "hyper-rules", which means that they take *derivations* as input. In this context, the notation '$\{\alpha_1;\ldots;\alpha_k \vdash \beta\}$' means that the lines in question form a derivation of β from $\alpha_1, \ldots, \alpha_k$. Accordingly, the notation '$\{\vdash\beta\}$' means the lines in question form a proof of β.

(r2) $\{\vdash \mathbb{T}\phi\} \hookrightarrow \Box\phi$

(r3) $\{\mathsf{FOL}\vdash\phi\} \hookrightarrow \Box\phi$

Here, $\{\mathsf{FOL}\vdash\phi\}$ is a set of earlier lines in the derivation that constitute a *first-order logic proof* of ϕ.[26]

12.2. Many of the Key Predicates of $\mathcal{T}$ are Necessitative

The next principles tell us that many of the key predicates of $\mathcal{T}$ are necessitative. Whether they are satisfied by an object (or pair of objects) is a matter of strict-necessity.

(a29) $\mathbb{S}x \rightarrow \Box\mathbb{S}x$
$\sim\mathbb{S}x \rightarrow \Box{\sim}\mathbb{S}x$

(a30) $\mathbb{P}x \rightarrow \Box\mathbb{P}x$
$\sim\mathbb{P}x \rightarrow \Box{\sim}\mathbb{P}x$

(a31) $\mathbb{N}x \rightarrow \Box\mathbb{N}x$
$\sim\mathbb{N}x \rightarrow \Box{\sim}\mathbb{N}x$

(a32) $x = y \rightarrow \Box\{x = y\}$
$x \neq y \rightarrow \Box\{x \neq y\}$

(a33) $x \precsim y \rightarrow \Box\{x \precsim y\}$
$\sim\{x \precsim y\} \rightarrow \Box{\sim}\{x \precsim y\}$

12.3. Optimization Principles

The following modal principles are added to ensure that the relation between the proposition $\langle\mathbb{T}\ulcorner\phi\urcorner\rangle$ and the proposition $\langle\phi\rangle$ is optimal.

(a34) $\mathbb{T}\ulcorner\phi\urcorner \leqslant \phi \;.\!\rightarrow\!.\; \mathbb{T}\ulcorner\phi\urcorner = \phi$

(a35) $\phi \leqslant \mathbb{T}\ulcorner\phi\urcorner \;.\!\rightarrow\!.\; \mathbb{T}\ulcorner\phi\urcorner = \phi$

12.4. Some Modal Theorems

We conclude this section with the following theorems.

(t35) $x \leqslant y \rightarrow \Box\{x \leqslant y\}$

(t36) $\sim\{x \leqslant y\} \rightarrow \Box{\sim}\{x \leqslant y\}$

---

[26] Note carefully that we cannot add the intuitively obvious modal principle,
$\{\vdash\phi\} \hookrightarrow \Box\phi$,
because it leads to inconsistency! See Section 26.1.

(t37) $x \equiv y \rightarrow \Box\{x \equiv y\}$

(t38) $\sim\{x \equiv y\} \rightarrow \Box\sim\{x \equiv y\}$

(t39) $x \approx y \rightarrow \Box\{x \approx y\}$

(t40) $\sim\{x \approx y\} \rightarrow \Box\sim\{x \approx y\}$

(t41) $\Box\phi \rightarrow \phi$

(t42) $\boxdot\phi \rightarrow \phi$

(t43) $\Box(\phi\rightarrow\psi) \rightarrow.\ \Box\phi \rightarrow \Box\psi$

(t44) $\boxdot(\phi\rightarrow\psi) \rightarrow.\ \boxdot\phi \rightarrow \boxdot\psi$

(t45) $\Box\phi \rightarrow \Box\Box\phi$

(t46) $\boxdot\phi \rightarrow \Box\boxdot\phi$

(t47) $\boxdot\phi \rightarrow \boxdot\boxdot\phi$

(t48) $\sim\Box\phi \rightarrow \Box\sim\Box\phi$

(t49) $\sim\boxdot\phi \rightarrow \Box\sim\boxdot\phi$

(t50) $\sim\boxdot\phi \rightarrow \boxdot\sim\boxdot\phi$

The reader familiar with modal logic will recognize most of these principles as principles of modal system S5.

13. '…Says…'

13.1. 'Says' is a Species of 'Implies'

We now turn to a key notion of theory $\mathcal{T}$ – 'says' – which is defined as follows.

(d19) ***x*** says ***y*** $=_{df}$ $\mathbb{S}\boldsymbol{x}\ \&\ \mathbb{P}\boldsymbol{y}\ \&\ \boldsymbol{x} \precsim \boldsymbol{y}$

In other words,

Sentence S *says* proposition p if and only if S *implies* p. Sentence S *says* that(ϕ) if and only if S *implies* that(ϕ).

In other words, 'says' is a species of 'implies' restricted so that only a sentence *can say*, and only a proposition *can be said.*[27]

The following are immediate theorems.

(t51) $s \text{ says } p \leftrightarrow s \precsim p$

(t52) $s \text{ says } p \rightarrow\!.\; p \precsim q \rightarrow s \text{ says } q$

(t53) $s \text{ says } p \;\&\; s \text{ says } q \;.\!\rightarrow s \text{ says } p{\wedge}q$

(t54) $s \text{ says } \langle\phi_1\rangle \;\&\; s \text{ says } \langle\phi_2\rangle \;.\!\rightarrow s \text{ says } \langle\phi_1 \,\&\, \phi_2\rangle$

(t55) $s \text{ says } \langle\phi_1\rangle \;\&\; s \text{ says } \langle\phi_1{\rightarrow}\phi_2\rangle \;.\!\rightarrow s \text{ says } \langle\phi_2\rangle$

(t56) $\forall\nu\{s \text{ says } \langle\phi\rangle\} \rightarrow s \text{ says } \langle\forall\nu\phi\rangle$

13.2. Motivating the Way $\mathcal{T}$ Formalizes 'says'.

By way of motivating our usage, we consider some examples from English. Suppose

I say that{snow is white **and** grass is green}

Question? Did I say that snow is white? Did I say that grass is green? I think the answer to both questions is unequivocally 'yes'.

Similarly, suppose

I say that{snow is white}
and
I say that{**if** snow is white, **then** grass is green}

Question? Did I say that grass is green? I think the answer, once again, is unequivocally 'yes'.

Finally, suppose

I say that{I am leaving **immediately**}

Question? Did I say that I am leaving. Once again, the answer is 'yes'.

This is exactly the way we propose to use the word 'says'. In particular, we treat "implicitly" saying as a primitive theoretical notion and "explicitly" saying [saying "in so many words"] as a purely pragmatic notion.[28]

---

[27] Note carefully that, although we only allow sentences and propositions to stand in the *logical* relation of saying, we do not rule out a *quasi-logical*[27] version of 'says', in the optional vocabulary, whose denotation could stand between people (or other ordinary objects) and propositions (and sentences).

[28] For example, one might ask 'did Max say that he is retiring?" You might answer 'yes, but not in so many words'. This means, I think, that Max *explicitly* said a number of things which together *imply* that he is retiring. If this is the situation, then Max *did* say that he was retiring, but *not in so many words*.

13.3. What a Sentence Says

Connected with the predicate ‘says’ is the *informal* function sign ‘what…says’. Grammatically, ‘what…says’ takes a sentence-denoting expression ***s*** and delivers a *noun phrase* – ‘what(***s*** says)’.[29]

This produces an immediate logico-grammatical problem. Given that a sentence ***s*** says (i.e., implies) *lots* of things, what(***s*** says) is not a single thing, but rather a bunch of things. Accordingly, ‘what(***s*** says)’ *is* a *plural* term, not a *singular* term. Given that first-order logic does not countenance irreducible plural terms, we have a problem. We solve this problem by using the expression ‘what(***s*** says)’ *informally*. For most purposes, the following will serve as a suggestive intermediate paraphrase.

(*id*) what(***s*** says) $\approx_{\text{df}}$ {x : ***s*** says x}

Here, {x : ***s*** says x} is the *plurality* (or, if you wish, set or class) of things that ***s*** says. Now, the specific formal rendering will depend upon the formula in which the expression appears. Let us do some examples.

Example 1:

s is true if and only if what(***s*** says) is true

Initial paraphrase:

s is true if and only if {x : ***s*** says x} is true

Let us understand that a plurality is true if and only if every element of that plurality is true, so this is equivalent to:

s is true if and only if everything(***s*** says) is true

The latter is formally rendered in $\mathcal{T}$ as follows.

s is true $\leftrightarrow$ $\forall$x{***s*** says x $\rightarrow$ x is true}

Example 2:

what($\boldsymbol{s_1}$ says) ***is*** what($\boldsymbol{s_2}$ says)

Initial paraphrase:

{x : $\boldsymbol{s_1}$ says x} = {x : $\boldsymbol{s_2}$ says x}

The latter is formally rendered in $\mathcal{T}$ as follows.

$\forall$x{$\boldsymbol{s_1}$ says x $\leftrightarrow$ $\boldsymbol{s_2}$ says x}

Example 3:

what($\boldsymbol{s_1}$ says) ***is*** p

---

[29] There is no generally accepted method of writing *out-fix* functors. The parentheses are inserted (admittedly awkwardly) precisely when they help visual parsing. For example, in ordinary English at least we must grammatically distinguish ‘what(S says) is true’, which is a sentence, from ‘what(S says is true)’, which is a noun phrase.

Here, things are a bit trickier. We need to re-render the singular ‘p’ as a corresponding plural. The natural plural counterpart of ‘p’ is ‘the propositions p implies’, in which case we have the following.

Initial paraphrase:

$$\{x : \boldsymbol{s_1} \text{ says } x\} = \{q : p \text{ implies } q\}$$

The latter is formally rendered in $\mathcal{T}$ as follows.

$$\forall q\{\boldsymbol{s_1} \text{ says } q \leftrightarrow p \precsim q\}$$

i.e.,

$$\forall q\{\boldsymbol{s_1} \precsim q \leftrightarrow p \precsim q\}$$

14. Standard Quotation Principles

We now turn to a crucial juncture for any theory of truth — how to deal with the following well-known quotation principles?

??	(q1)	that(ϕ) is true iff ϕ
??	(q2)	$\ulcorner\phi\urcorner$ is true iff ϕ
??	(q3)	$\ulcorner\phi\urcorner$ is true iff that(ϕ) is true

First, we divide these principles into their natural parts and name them.

(q1)	$\mathbb{T}\langle\phi\rangle \leftrightarrow \phi$		[indirect quotation schema]
(a)	$\mathbb{T}\langle\phi\rangle \rightarrow \phi$		[indirect **dis**-quotation]
(b)	$\phi \rightarrow \mathbb{T}\langle\phi\rangle$		[indirect **en**-quotation]
(q2)	$\mathbb{T}\ulcorner\phi\urcorner \leftrightarrow \phi$		[direct quotation schema]
(a)	$\mathbb{T}\ulcorner\phi\urcorner \rightarrow \phi$		[direct **dis**-quotation]
(b)	$\phi \rightarrow \mathbb{T}\ulcorner\phi\urcorner$		[direct **en**-quotation]
(q3)	$\mathbb{T}\ulcorner\phi\urcorner \leftrightarrow \mathbb{T}\langle\phi\rangle$		[the “denotation” principle][30]
(a)	$\mathbb{T}\ulcorner\phi\urcorner \rightarrow \mathbb{T}\langle\phi\rangle$		[direct **dis**-quotation + indirect-**en**-quotation]
(b)	$\mathbb{T}\langle\phi\rangle \rightarrow \mathbb{T}\ulcorner\phi\urcorner$		[indirect **dis**-quotation + direct **en**-quotation]

15. The Naive Theory of Truth

We next briefly consider adding (q1)–(q3) to system $\mathcal{T}$. The resulting system represents our formulation of the Naive Theory of Truth (NTT). As we will see in the next section, NTT is problematic at best, so we will have to make some hard decisions concerning which components of (q1)–(q3) to adopt.

16. The Liar Paradox

Having explicitly formulated NTT, we now show what exactly is problematic about it. This brings us to the most famous counterexample in the history of philosophy, the Liar Paradox, proposed originally by Eubulides.[31] Consider the sentence.

(λ) this sentence is not true

In and of itself, λ is not problematic. The demonstrative ‘this sentence’ can presumably be used to refer to *any* sentence. For example, it can be used to refer to

(r) 2+2=5

in which case λ is unproblematically true.

The problem arises when the demonstrative ‘this sentence’ in λ is used to refer to λ *itself* — in other words, when we explicitly add the following further claim.

($\lambda^=$) this sentence = ‘this sentence is not true’

[30] This is called the “denotation principle” because it roughly says that a sentence is true if and only if its denotation is true. Here, the term ‘denotation’ is purely heuristic, being borrowed from the meta-theory. In particular, according to the semantics for $\mathcal{T}$, every admissible valuation υ will assign to each formula of $\mathcal{L}$ – and to each proposition-denoting noun-phrase – a “proposition”. Furthermore, for each formula of $\mathcal{L}$, $\upsilon(\phi) = \upsilon(\langle\phi\rangle)$.

[31] Eubulides of Megara proposed the Liar paradox as a counter-example to Aristotle’s hugely famous definition of truth. Some people *nominally* attribute the Liar Paradox to Epimenides. This is unfortunate, since it suggests that Epimenides formulated it. But he did not *formulate* the Liar Paradox; he was merely an *unwitting victim* of it.

Before continuing, however, we first translate this sentence into the formal language of $\mathcal{T}$.

($\lambda^{=}$) $\lambda = \ulcorner\lambda$ is not true$\urcorner$

We are now in a position to produce some serious mischief! First, applying the direct-quotation schema

(q1) $\ulcorner\phi\urcorner$ is true $\leftrightarrow \phi$

we obtain:

$\ulcorner\lambda$ is not true$\urcorner$ is true $\leftrightarrow$ λ is not true

Then applying Leibniz's Law, in connection with ($\lambda^{=}$), we obtain:

λ is true $\leftrightarrow$ λ is *not* true

This is of course a self-contradiction!

Notice carefully that we have *not* shown that NTT is logically inconsistent *per se*, because ($\lambda^{=}$) is, after all, a *supposition*. All we have shown is that NTT is inconsistent with ($\lambda^{=}$). Alternatively stated, we have not deduced a *contradiction* from NTT; we have deduced a *theorem schema*! Namely:

(t!) $\boldsymbol{s} \neq \ulcorner\boldsymbol{s}$ is not true$\urcorner$[32]

Unfortunately, however, allowing (t!) as a theorem wreaks havoc in the meta-theory. In particular, when we do formal semantics, we stipulate that any given variable ***x*** *can* denote *any* item in the universe of discourse; otherwise, it is impossible to explicate quantification *over* ***x***. This applies *mutatis mutandis* to sortal variables. In particular, we stipulate that any given sentence-denoting variable ***s*** *can* denote *any* item in *its* domain.

Now, the intended domain of "sentences" is the set of all formulas of the language of $\mathcal{T}$.[33] This means that it is *semantically permissible* for the denotation of ***s*** to be the formula [$\sim\mathbb{T}\boldsymbol{s}$]. It is accordingly *semantically permissible* for [$\boldsymbol{s} = \ulcorner\sim\mathbb{T}\boldsymbol{s}\urcorner$] to be true. Accordingly, [$\boldsymbol{s} \neq \ulcorner\sim\mathbb{T}\boldsymbol{s}\urcorner$] cannot be a theorem of any logical system that does quantification in the usual way (i.e., by allowing every variable to "vary" over its domain).

There is a natural language counterpart of the reasoning in the previous paragraph, given by our earlier sentence.

(λ) this sentence is not true

[32] Note carefully that there is no corresponding *single* theorem of this "type". For example, the following could not be a theorem of any consistent theory:

$\forall s\{s \neq \ulcorner s$ is not true$\urcorner\}$,

since it has the following as an immediate logical consequence!

$\ulcorner s$ is not true$\urcorner \neq \ulcorner s$ is not true$\urcorner$

[33] More generally, we have an arbitrary $\mathcal{T}$-language; see Section (24).

Now, just as a variable ***x*** can refer to any item in its domain, the demonstrative 'this sentence' can presumably be used to denote *any* sentence. So in particular, it can be used to denote 'this sentence is not true'. In other words, the following is *semantically admissible*.

($\lambda^{=}$) this sentence = 'this sentence is not true'

17. Damage Control

As stated at the beginning, our goal is to save as much of the Naive Theory of Truth as we can. In pursuit of this goal, we need to pinpoint the error that produces the Liar Paradox. Toward that end we make a few observations.

First, any two of the standard quotation principles – (q1)–(q3) – imply the third; that means that we cannot abandon *just one* of them. On the other hand, no singleton implies either of the other two; that means that we can maintain one of the quotation principles, while abandoning the other two. Finally, we observe that we can reject half of a biconditional without rejecting the other half.

In other words, there are *in principle* a number of logical choices concerning which specific components of (q1)–(q3) to maintain in one's theory of truth. Rather than discuss the possibilities, we simply declare our own choices.

17.1. The Indirect-Quotation Principle

First, we declare the following.

$\mathcal{T}$ accepts **both** the indirect-**dis**-quotation schema [that(ϕ) is true $\rightarrow$ ϕ] **and** the indirect-**en**-quotation schema [ϕ $\rightarrow$ that(ϕ) is true]

We make this official by adding the following theorem schemata.

(t57) $\mathbb{T}\langle\phi\rangle \leftrightarrow \phi$ [indirect quotation schema]

(t58) $\mathbb{T}\langle\phi\rangle \rightarrow \phi$ [indirect **dis**-quotation]

(t59) $\phi \rightarrow \mathbb{T}\langle\phi\rangle$ [indirect **en**-quotation]

Note, these are not stated as axioms. That is because, we posit the following considerably stronger axiom, which is followed by the obvious corollary.

(a36) $\langle\mathbb{T}p\rangle = p$

(t60) $\langle\mathbb{T}\langle\phi\rangle\rangle = \langle\phi\rangle$

In other words,

The proposition that(ϕ)
and the proposition that(it is true that ϕ)[34]
are the very same proposition.

17.2. Direct-Dis-Quotation

Next, we declare the following.

$\mathcal{T}$ accepts the direct-**dis**-quotation schema:
$\ulcorner\phi\urcorner$ is true $\rightarrow$ ϕ

This is officially implemented in $\mathcal{T}$ by the following axiom schema.

(a37) $\mathbb{T}\ulcorner\phi\urcorner \rightarrow \phi$ [direct **dis**-quotation]

Notice that, in light of earlier results about 'implies' and 'says', this is *equivalent* to the following rather innocuous result.

(t61) $\ulcorner\phi\urcorner$ says $\langle\phi\rangle$

In other words,

The sentence $\ulcorner\phi\urcorner$ says that(ϕ).

17.3. What We Must Reject

Now, if we accept (a37), then we must reject its converse,

✕ $\phi \rightarrow \mathbb{T}\ulcorner\phi\urcorner$ [✕direct **en**-quotation✕]

because together they produce the Liar Paradox. Also, with (t57) and (a37) in hand, we can prove the following theorem.

(t62) $\mathbb{T}\ulcorner\phi\urcorner \rightarrow \mathbb{T}\langle\phi\rangle$ [direct **dis**-quotation + indirect **en**-quotation]

But we must reject its converse,

✕ $\mathbb{T}\langle\phi\rangle \rightarrow \mathbb{T}\ulcorner\phi\urcorner$ [✕indirect **dis**-quotation + direct **en**-quotation✕]

since together with (t57) and (t62) it produces the Liar Paradox.

[34] Recall that we take 'it is true that ϕ' and 'that ϕ is true' simply to be grammatical variants of each other. See Section 9.1.

18. Summary of Quotation Principles of $\mathcal{T}$

When all is said and done, we have the following quotation schemata.

(q1) $\mathbb{T}\langle\phi\rangle \leftrightarrow \phi$ [indirect **dis**-quotation and **en**-quotation]

(q2a) $\mathbb{T}\ulcorner\phi\urcorner \rightarrow \phi$ [direct **dis**-quotation]

(q3a) $\mathbb{T}\ulcorner\phi\urcorner \rightarrow \mathbb{T}\langle\phi\rangle$ [direct **dis**-quotation + indirect **en**-quotation]

But we reject the following schemata.

✕ $\phi \rightarrow \mathbb{T}\ulcorner\phi\urcorner$ [✕direct **en**-quotation✕]

✕ $\mathbb{T}\langle\phi\rangle \rightarrow \mathbb{T}\ulcorner\phi\urcorner$ [✕indirect **dis**-quotation + direct **en**-quotation✕]

Note carefully, however, that ultimately we only reject a *tiny* minority of applications of these two principles. As we see later, most applications of direct-en-quotation (the useful ones!) can be salvaged. [See Section 28.]

19. The Fundamental Truth-Principle

Having officially stated the quotation principles of System $\mathcal{T}$, we can now succinctly state the fundamental principle of $\mathcal{T}$ about the truth of sentences.

(fpt) $\mathbb{T}s \leftrightarrow \forall x\{s \text{ says } x \rightarrow \mathbb{T}x\}$

A sentence is true iff everything(it says) is true.

First note that we can use our *informal* notion of 'what(***s*** says)' (recall Section 13.3) to rephrase the fundamental truth-principle as follows.

A sentence is true iff what(it says) is true.

In order to prove the fundamental truth-principle, we first prove an important lemma.

(t63) S says that(S is true)

Every sentence says of itself that it is true.

Even though the proof is quite simple, we include it since the theorem may seem unintuitive.[35] It comes down to the following facts about $\mathcal{T}$.

(1) ‘says’ means ‘implies’;
(2) ‘s implies *x*’ means ‘that(s is true) implies that(*x* is true)’;
(3) indirect en-quotation is valid in $\mathcal{T}$.

(1)	$\mathbb{T}s \rightarrow \mathbb{T}\langle\mathbb{T}s\rangle$	indirect-en-quotation
(2)	$\boxdot\{\mathbb{T}s \rightarrow \mathbb{T}\langle\mathbb{T}s\rangle\}$	1(⊢),(r1)
(3)	$x \precsim y \leftrightarrow \boxdot(\mathbb{T}x \rightarrow \mathbb{T}y)$	(a27)
(4)	$s \precsim \langle\mathbb{T}s\rangle \leftrightarrow \boxdot\{\mathbb{T}s \rightarrow \mathbb{T}\langle\mathbb{T}s\rangle\}$	instance of 3
(5)	$s \precsim \langle\mathbb{T}s\rangle$	2,4,SL
(6)	s says $\langle\mathbb{T}s\rangle$	5,Def(says)

Now with our key lemma in hand, we can prove the fundamental truth-principle.

(t64) $\mathbb{T}s \leftrightarrow \forall x\{s \text{ says } x \rightarrow \mathbb{T}x\}$

	[→]	
(1)	$\mathbb{T}s$	As
(2)	~~SHOW~~: $\forall x\{s \text{ says } x \rightarrow \mathbb{T}x\}$	UCD
(3)	│ s says p	As
(4)	│ ~~SHOW~~: $\mathbb{T}p$	DD
(5)	│ │ $s \precsim p$	3,Def(says)
(6)	│ │ $\mathbb{T}p$	1,5,≾-OUT
	[←]	
(7)	$\forall x\{s \text{ says } x \rightarrow \mathbb{T}x\}$	As
(8)	~~SHOW~~: $\mathbb{T}s$	DD
(9)	│ s says $\langle\mathbb{T}s\rangle$	(t63)
(10)	│ $\mathbb{T}\langle\mathbb{T}s\rangle$	7,9,QL
(11)	│ $\mathbb{T}s$	10,indirect-**dis**-quotation

20. How the Fundamental Truth-Principle *Sometimes* Delivers the Naive Direct-Quotation Principle

We are now in position to state a fundamental theorem about the naive direct-quotation principle.

(t65) $\forall x\{\ulcorner\phi\urcorner \text{ says } x \rightarrow \langle\phi\rangle \precsim x\} \rightarrow\!. \; \mathbb{T}\ulcorner\phi\urcorner \leftrightarrow \phi$

In other words,

$\mathbb{T}\ulcorner\phi\urcorner \leftrightarrow \phi$ ***provided*** $\langle\phi\rangle$ implies everything $\ulcorner\phi\urcorner$ says
$\mathbb{T}\ulcorner\phi\urcorner \leftrightarrow \phi$ ***provided*** $\langle\phi\rangle$ implies what($\ulcorner\phi\urcorner$ says)

[35] The formal proof technique employed here is an expansion of a system is presented in Hardegree, *Symbolic Logic: A First Course* (McGraw-Hill). It is expanded to include numerous “obvious” short-cuts, including reasoning by SL (sentential logic), QL (quantifier logic), and IL (identity logic).

Now, the basic idea is that *ordinarily*, the proposition $\langle\phi\rangle$ does imply everything the sentence $\ulcorner\phi\urcorner$ says. But unfortunately, this is not universally true, given the Liar Paradox.

(1)	$\forall x\{\ulcorner\phi\urcorner$ says $x \rightarrow \langle\phi\rangle \precsim x\}$	As
(2)	~~SHOW~~: $\mathbb{T}\ulcorner\phi\urcorner \leftrightarrow \phi$	3,4,SL
(3)	$[\rightarrow]$	direct-**dis**-quotation
(4)	$[\leftarrow]$	CD
(5)	| ϕ	As
(6)	| ~~SHOW~~: $\mathbb{T}\ulcorner\phi\urcorner$	7+(t64)
(7)	| ~~SHOW~~: $\forall x\{\ulcorner\phi\urcorner$ says $x \rightarrow \mathbb{T}x\}$	UCD
(8)	| | $\ulcorner\phi\urcorner$ says p	As
(9)	| | ~~SHOW~~: $\mathbb{T}p$	DD
(10)	| | | $\langle\phi\rangle \precsim p$	1,8,QL
(11)	| | | $\mathbb{T}\langle\phi\rangle$	5,indirect-en-quotation
(12)	| | | $\mathbb{T}p$	10,11,$\precsim$-OUT

21. The Liar Reconsidered

Having explicitly formulated the fundamental truth-principles for $\mathcal{T}$, we now reconsider the liar sentence, which is formalized in $\mathcal{T}$ as follows.

($\lambda^=$) $\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner$

First, we show the following important theorem, which is followed by two even more important *non*-theorems.

(t66) $\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner \rightarrow {\sim}\mathbb{T}\lambda$

> The liar sentence λ is not true.[36]

(1)	$\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner$	As
(2)	~~SHOW~~: ${\sim}\mathbb{T}\lambda$	ID
(3)	| $\mathbb{T}\lambda$	As
(4)	| ~~SHOW~~: ✕	3,6,SL
(5)	| | $\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner$	1,3,IL
(6)	| | ${\sim}\mathbb{T}\lambda$	5+direct-**dis**-quotation (a37)

What are *not* theorems are the following.[37]

✕ $\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner \rightarrow.\ \mathbb{T}\lambda\ \&\ {\sim}\mathbb{T}\lambda$

✕ $\lambda \neq \ulcorner\lambda$ is not true$\urcorner$

For example, the following reasoning is now *officially banned* in $\mathcal{T}$.

[36] The expression 'the liar sentence' is misleading. In the official formal system, there are no liar sentences, but only liar situations. A *liar-situation* is a situation in which (for some s), s = $\ulcorner\ \mathbb{T}s\urcorner$. Thus, when we use the expression "the liar sentence λ is such that $\Phi\lambda$", we really mean "if $\lambda = \ulcorner\ \mathbb{T}\lambda\urcorner$, then $\Phi\lambda$".

[37] As expected, this *non–theorem* is contingent upon an appropriate consistency proof for $\mathcal{T}$.

(7)	$\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner$	✕✕3, direct-**en**-quotation✕✕
(8)	$\mathbb{T}\lambda$	1,7,IL
(9)	✕	2,8,SL

Usually, we disguise the invalid step by reasoning *informally* as follows.

✕ We have already shown that λ is not true (line 2). But *that* is **exactly** what λ says! So λ is true. Thus, λ is true, **and** λ is not true, which is a contradiction!

22. The Critical Juncture!

22.1. Is *That* Exactly What λ Says?

We have now reached the critical juncture in the development of our logic of truth. Consider the informal argument immediately above. We agree with all the reasoning **except** the step:

✕ *that* is **exactly** what λ says!

In other words,

✕ that(λ is not true) is exactly what(λ says)

We agree that if the previous step is correct, then the reasoning proceeds validly. So we need to explain why the previous step is incorrect.

22.2. Formalizing '*that* is exactly what s says'

In formalizing the questionable statement, we must be very careful. In particular, recall that 'what…says' has a purely *informal* status in $\mathcal{T}$ (Section 13.3), the following being the rough-and-ready paraphrase scheme.

$$\text{what}(\boldsymbol{s}\text{ says}) \quad \approx_{\text{df}} \quad \{x : \boldsymbol{s}\text{ says } x\}$$

The latter is understood as a "plural" that must itself be paraphrased away formally. In the case of an informal formula

$$\text{p is exactly what } \boldsymbol{s} \text{ says}$$

we begin by rewriting it as

$$\text{p} = \text{what}(\boldsymbol{s}\text{ says})$$

$$\text{p} = \{x : \boldsymbol{s}\text{ says } x\}$$

Then we the rewrite 'p' as a plural 'the propositions p implies' to obtain

$$\{q\text{: p} \precsim q\} = \{x : \boldsymbol{s}\text{ says } x\}$$

the latter being formally rendered in $\mathcal{T}$ as follows.

$\forall$q{***s*** says q $\leftrightarrow$ p $\precsim$ q}

22.3. An Alternative Formalization Using the Notion '*and nothing more*'

An alternative route to formally rendering 'that is exactly what ***s*** says' employs the notion 'and nothing more'. In particular, our intermediate paraphrase is:

p is exactly what ***s*** says $=_{df}$ ***s*** says p *and nothing more*

The latter is notion is filled out by the following definition.

(d20) ***s*** says ***p*** *and nothing more* $=_{df}$ ***s*** says ***p*** & $\forall$x{***s*** says x $\rightarrow$ ***p*** $\precsim$ x}

In other words,

s says **p** *and nothing more* means: **s** says **p** , **and p** implies everything(**s** says).

We next observe that the definition could be logically simplified, in light of the following theorem.

(t67) s says p *and nothing more* $\leftrightarrow$ $\forall$x{s says x $\leftrightarrow$ p $\precsim$ x}

One can now immediately see that this formalization of 'p is exactly what s says' is in complete agreement with the one in the previous section.

22.4. Suppose p is exactly what S says [i.e., S says p *and nothing more*]

Also, with this definition in hand, we can show that, granting the questionable step 'that is exactly what λ says', permits the argument to proceed to it usual paradoxical conclusion. In particular, we have the following theorems. In this connection, compare these results with the theorem in Section 20, which anticipates them.

(t68) S says p *and nothing more* $\rightarrow$. S is true $\leftrightarrow$ p is true

(t69) S says $\langle\phi\rangle$ *and nothing more* $\rightarrow$. S is true $\leftrightarrow$ ϕ

(t70) $\ulcorner\phi\urcorner$ says $\langle\phi\rangle$ *and nothing more* $\rightarrow$. $\ulcorner\phi\urcorner$ is true $\leftrightarrow$ ϕ

(t71) λ says $\langle\sim\mathbb{T}\lambda\rangle$ *and nothing more* $\rightarrow$. $\mathbb{T}\lambda$ $\leftrightarrow$ $\sim\mathbb{T}\lambda$

22.5. Back to the Original Question

Given the theorems at the end of the previous section, we must now ask whether

? λ says $\langle\sim\mathbb{T}\lambda\rangle$ *and nothing more*

i.e.,

? λ says $\langle \sim\mathbb{T}\lambda \rangle$, **and** $\langle \sim\mathbb{T}\lambda \rangle$ implies everything(λ says)

We agree that since

$$\lambda = \ulcorner \sim\mathbb{T}\lambda \urcorner$$

it follows that

λ says $\langle \sim\mathbb{T}\lambda \rangle$ [i.e., $\lambda \precsim \langle \sim\mathbb{T}\lambda \rangle$]

since this is a simple instance of an earlier theorem schema (t69).

$\ulcorner \phi \urcorner$ says $\langle \phi \rangle$

But we do not agree that

✕ $\langle \sim\mathbb{T}\lambda \rangle$ implies everything(λ says)

In particular, from

λ says $\langle \sim\mathbb{T}\lambda \rangle$

we *cannot* deduce that

✕ λ does *not* say anything *else*,

or:

✕ λ says $\langle \sim\mathbb{T}\lambda \rangle$ *and nothing more*,

or:

✕ $\langle \sim\mathbb{T}\lambda \rangle$ is exactly what(λ says).

22.6. Our Version of the Liar Paradox

Indeed, we are now able to state exactly what *is* "paradoxical" about the liar sentence λ.

> Although the liar sentence λ
> says of itself that it is not true,
> no sentence, including λ,
> can say of itself that it is not true *and nothing more*.

(t72) $\lambda = \ulcorner \sim\mathbb{T}\lambda \urcorner \rightarrow \lambda$ says $\langle \sim\mathbb{T}\lambda \rangle$

This is an immediate consequence of (t61).

(t73) $\sim\exists$s[s says that s is not true <u>and nothing more</u>]

This is an immediate consequence of (t69).

22.7. The Definition of "a Liar Sentence"

So, suppose someone proposes to *define* a "liar sentence" as follows.

(d?) a "liar sentence" is, by *definition*, a sentence
that says of itself that it is not true *and nothing else* (and hence *nothing more*)

Then our response is quite simple. There is no such thing, since its existence is logically impossible! Suppose that the same person responds by pointing to the sentence

this sentence is not true.

Then our response is equally simple. Although you have exhibited a sentence which *manifestly* says of itself that it is not true, you have not exhibited a sentence that *manifestly* says *only* that it is not true.[38]

[38] In this connection, it is important to note that, although Tarski proved the existence of liar sentences in RA+T, he did not prove the existence of sentences that say of themselves *only* that they are not true.

23. What *Else* Does The Liar Say?

We have shown that, although the liar sentence says of itself that it is not true, *that* is not *all* that the liar sentence says. But what *else does* it say? That is easy. In fact, we already answered this question in an earlier section. For, you see, we have already established the following theorem .

(t63) S says that(S is true)

> *Every* sentence says of itself that it is true.

Combining this with the theorem that λ says that λ is not true, we obtain the following theorems.[39]

(t74) λ says $\langle\lambda$ is not true$\rangle$ & λ says $\langle\lambda$ is true$\rangle$

(t75) λ says $\langle\lambda$ is not true & λ is true$\rangle$

(t76) λ says the impossible proposition

(t77) $\boxdot\sim\mathbb{T}\lambda$

> The liar sentence says of itself that it **is not** true,
> **but** (as with every sentence) it also says of itself that it **is** true.
> and so:
> The liar sentence cannot be true!

24. The Deferred Liar Paradox

As most people are aware, there are other liar-like paradoxes. We will examine several of these over the course of this paper, in step with our development of the necessary formal apparatus. We are now in a position to consider the "deferred liar" (also called "extended liar" and "liar cycle"), which may be stated colloquially as follows.

(1) the following sentence is true;
(2) the previous sentence is not true.

This example is the two-sentence version of the deferred liar. In fact, for any natural number n, there is an n-sentence version of the deferred liar. However, there are no further insights gained in the general case, so we will simply concentrate on the two-sentence deferred liar.

First, we reformulate the sentences in the formal language of $\mathcal{T}$ as follows.

(s1) $\lambda_1 = \ulcorner\mathbb{T}\lambda_2\urcorner$
(s2) $\lambda_2 = \ulcorner\sim\mathbb{T}\lambda_1\urcorner$

---

[39] In regard to the reference of 'λ' and 'the liar sentence', please see footnote 36.

Next, we show that these two suppositions produce a contradiction when we submit them to the naive truth-schema (direct quotation principle).

(1)	$\lambda_1 = \ulcorner \mathbb{T}\lambda_2 \urcorner$	Pr
(2)	$s_2 = \ulcorner {\sim}\mathbb{T}\lambda_1 \urcorner$	Pr
(3)	~~SHOW~~: ✕	8,SL
(4)	$\mathbb{T}\ulcorner \mathbb{T}\lambda_2 \urcorner \leftrightarrow \mathbb{T}\lambda_2$	✕Direct-Quotation Principle✕
(5)	$\mathbb{T}\lambda_1 \leftrightarrow \mathbb{T}\lambda_2$	1,4,IL
(6)	$\mathbb{T}\ulcorner {\sim}\mathbb{T}\lambda_1 \urcorner \leftrightarrow {\sim}\mathbb{T}\lambda_1$	✕Direct-Quotation Principle✕
(7)	$\mathbb{T}\lambda_2 \leftrightarrow {\sim}\mathbb{T}\lambda_1$	2,5,IL
(8)	$\mathbb{T}\lambda_1 \leftrightarrow {\sim}\mathbb{T}\lambda_1$	5,7,SL

Notice, of course, that this "proof" has two steps that appeal to the direct-quotation principle, which we have officially banished.

Next, we construct a genuine proof in $\mathcal{T}$, showing the following theorem.

(t78) $\lambda_1 = \ulcorner \mathbb{T}\lambda_2 \urcorner \ \& \ \lambda_2 = \ulcorner {\sim}\mathbb{T}\lambda_1 \urcorner \ .\!\rightarrow {\sim}\mathbb{T}\lambda_1$

(1)	$\lambda_1 = \ulcorner \mathbb{T}\lambda_2 \urcorner$	As
(2)	$\lambda_2 = \ulcorner {\sim}\mathbb{T}\lambda_1 \urcorner$	As
(3)	~~SHOW~~: ${\sim}\mathbb{T}\lambda_1$	ID
(4)	$\mathbb{T}\lambda_1$	As
(5)	~~SHOW~~: ✕	4,9,SL
(6)	$\mathbb{T}\ulcorner \mathbb{T}\lambda_2 \urcorner$	1,4,IL
(7)	$\mathbb{T}\lambda_2$	6,direct-dis-quotation
(8)	$\mathbb{T}\ulcorner {\sim}\mathbb{T}\lambda_1 \urcorner$	2,7,IL
(9)	${\sim}\mathbb{T}\lambda_1$	8,direct-dis-quotation

At this point, we will be sorely tempted to continue reasoning *informally* as follows.

✕ We have shown that λ_1 is not true (line 4). But *that* is **exactly** what λ_2 says! So λ_2 is true. But *that* is **exactly** what λ_1 says! So, λ_1 is true. But this contradicts our earlier line that λ_1 is not true!

Suppose that we formally write this argument down, line by line. Then it goes as follows.

(10)	$\mathbb{T}\ulcorner {\sim}\mathbb{T}\lambda_1 \urcorner$	3, ✕direct-**en**-quotation✕
(11)	$\mathbb{T}\lambda_2$	2,10,IL
(12)	$\mathbb{T}\ulcorner \mathbb{T}\lambda_2 \urcorner$	11, ✕direct-**en**-quotation✕
(13)	$\mathbb{T}\lambda_1$	1,12,IL

You will notice of course that this derivation involves two applications of direct-**en**-quotation, which we have officially banished!

25. A Further Truth-Principle

What can we say about the second sentence λ_2 in the deferred liar? At the moment, its truth-value is under-determined by our principles. We take this opportunity to correct this situation.

Specifically, we add the following *special instance* of the direct-**en**-quotation principle. [Note, it is presented as a theorem, not an axiom, because we generate it from a “later” *stronger* axiom.]

(t79) $\mathbb{T}s \rightarrow \mathbb{T}\ulcorner\mathbb{T}s\urcorner$

Notice that, given this new theorem, line (12) in the above derivation is now officially admitted – but, the derivation still has an invalid step at line (10)!

Given the general validity of direct-**dis**-quotation, we have the following immediate theorem.

(t80) $\mathbb{T}\ulcorner\mathbb{T}s\urcorner \leftrightarrow \mathbb{T}s$

Indeed, we considerably strengthen this by positing the following axiom.

(a38) $\langle\mathbb{T}\ulcorner\mathbb{T}s\urcorner\rangle = \langle\mathbb{T}s\rangle$

Note that we have the following as an immediate and useful theorem schema.

(t81) $\mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\urcorner\urcorner = \mathbb{T}\ulcorner\phi\urcorner$

With this axiom and resulting theorems in hand, we can now answer the question posed at the beginning of this section; in particular, the sentence in question, λ_2, is not true.

(t82) $\lambda_1 = \ulcorner\mathbb{T}\lambda_2\urcorner \;\&\; \lambda_2 = \ulcorner\sim\mathbb{T}\lambda_1\urcorner \;.\!\rightarrow \sim\mathbb{T}\lambda_2$

(1)	$\lambda_1 = \ulcorner\mathbb{T}\lambda_2\urcorner$	As
(2)	$\lambda_2 = \ulcorner\sim\mathbb{T}\lambda_1\urcorner$	As
(3)	~~SHOW~~: $\sim\mathbb{T}\lambda_2$	ID
(4)	│ $\mathbb{T}\lambda_2$	As
(5)	│ ~~SHOW~~: ✗	7,8,SL
(6)	│ │ $\mathbb{T}\ulcorner\mathbb{T}\lambda_2\urcorner$	4+(t79)
(7)	│ │ $\mathbb{T}\lambda_1$	1,6,IL
(8)	│ │ $\sim\mathbb{T}\lambda_1$	1,2+(t78)

26. Quasi-Modal Principles of Truth

We have so far obtained two direct-quotation principles, given as follows.

$\mathbb{T}\ulcorner\phi\urcorner \rightarrow \phi$

$\mathbb{T}\ulcorner\phi\urcorner \rightarrow \mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\urcorner\urcorner$

At this point, we observe that these two principles are *analogous* to the following modal principles.

(T) $\Box\phi \rightarrow \phi$

(4) $\Box\phi \rightarrow \Box\Box\phi$

These are both principles of modal system S4. As an exercise in “reverse engineering”, we next consider other quasi-modal alethic principles, to see if they are valid. In the following list, we write each principle in its pure modal form, and in its corresponding alethic form.[40] We also include the previous two principles, for the sake of thoroughness.

	Modal Principle	**Alethic Counterpart**
N	$\{\vdash\phi\} \hookrightarrow \Box\phi$	$\{\vdash\phi\} \hookrightarrow \mathbb{T}\ulcorner\phi\urcorner$
K	$\Box(\phi \to \psi) \to\!. \ \Box\phi \to \Box\psi$	$\mathbb{T}\ulcorner\phi \to \psi\urcorner \to\!. \ \mathbb{T}\ulcorner\phi\urcorner \to \mathbb{T}\ulcorner\psi\urcorner$
D	$\Box\phi \to \sim\Box\sim\phi$	$\mathbb{T}\ulcorner\phi\urcorner \to \sim\mathbb{T}\ulcorner\sim\phi\urcorner$
T	$\Box\phi \to \phi$	$\mathbb{T}\ulcorner\phi\urcorner \to \phi$
B	$\sim\Box\sim\Box\phi \to \phi$	$\sim\mathbb{T}\ulcorner\sim\mathbb{T}\ulcorner\phi\urcorner\urcorner \to \phi$
4	$\Box\phi \to \Box\Box\phi$	$\mathbb{T}\ulcorner\phi\urcorner \to \mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\urcorner\urcorner$
5	$\sim\Box\sim\Box\phi \to \Box\phi$	$\sim\mathbb{T}\ulcorner\sim\mathbb{T}\ulcorner\phi\urcorner\urcorner \to \mathbb{T}\ulcorner\phi\urcorner$
G	$\sim\Box\sim\Box\phi \to \Box\sim\Box\sim\phi$	$\sim\mathbb{T}\ulcorner\sim\mathbb{T}\ulcorner\phi\urcorner\urcorner \to \mathbb{T}\ulcorner\sim\mathbb{T}\ulcorner\sim\phi\urcorner\urcorner$
L	$\Box(\Box\phi \to \psi) \vee \Box(\Box\psi \to \phi)$	$\mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\urcorner \to \psi\urcorner \vee \mathbb{T}\ulcorner\mathbb{T}\ulcorner\psi\urcorner \to \phi\urcorner$
–5	$\Box\sim\Box\sim\phi \to \sim\Box\sim\Box\phi$	$\mathbb{T}\ulcorner\sim\mathbb{T}\sim\ulcorner\phi\urcorner\urcorner \to \sim\mathbb{T}\ulcorner\sim\mathbb{T}\ulcorner\phi\urcorner\urcorner$
–B	$\Box\sim\phi \to \Box\sim\Box\phi$	$\mathbb{T}\ulcorner\sim\phi\urcorner \to \mathbb{T}\ulcorner\sim\mathbb{T}\ulcorner\phi\urcorner\urcorner$
B^-	$\Box\sim\Box\sim\Box\phi \to \Box\phi$	$\mathbb{T}\ulcorner\sim\mathbb{T}\ulcorner\sim\mathbb{T}\ulcorner\phi\urcorner\urcorner\urcorner \to \mathbb{T}\ulcorner\phi\urcorner$
C	$\sim\Box\sim\phi \to \Box\phi$	$\sim\mathbb{T}\ulcorner\sim\phi\urcorner \to \mathbb{T}\ulcorner\phi\urcorner$
bf	$\Box\forall\nu\phi \leftrightarrow \forall\nu\Box\phi$	$\mathbb{T}\ulcorner\forall\nu\phi\urcorner \leftrightarrow \forall\nu\mathbb{T}\ulcorner\phi\urcorner$

As one would expect, some of these modal principles can be successfully transferred to System $\mathcal{T}$, and others cannot. We consider them in turn.

26.1. The Rule of Necessitation (Alethiation)

The alethic counterpart of the rule of necessitation (N) – what we will call the "rule of alethiation" – is the following rule.

✕ $\{\vdash\phi\} \hookrightarrow \mathbb{T}\ulcorner\phi\urcorner$

This rule *cannot* be added to System $\mathcal{T}$. For then we would have the following meta-theorem.

[40] Notice that there is no point in considering the corresponding alethic operator 'it is true *that*…', since we have already shown that this operator is redundant. In particular, $\mathbb{T}\langle\phi\rangle = \phi$

✕ $\vdash\phi \rightarrow \vdash\mathbb{T}\ulcorner\phi\urcorner$

To see that this is unacceptable, we recall that the direct-dis-quotation principle,

$$\mathbb{T}\ulcorner\phi\urcorner \rightarrow \phi$$

is a thesis of $\mathcal{T}$, so if general alethiation were admissible, the following would also be a thesis of $\mathcal{T}$

✕ $\mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\urcorner \rightarrow \phi\urcorner$

This does not look bad [in fact, it looks kind of obvious!] until we consider the liar sentence. In particular, consider the following instance.

✕ $\mathbb{T}\ulcorner\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner \rightarrow {\sim}\mathbb{T}\lambda\urcorner$

The following proof shows that this principle is inconsistent with the existence of a liar sentence.

(1)	$\mathbb{T}\ulcorner\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner \rightarrow {\sim}\mathbb{T}\lambda\urcorner$	Pr
(2)	$\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner$	Pr
(3)	~~SHOW~~: ✕	8,9,SL
(4)	$\mathbb{T}\lambda \rightarrow {\sim}\mathbb{T}\lambda \;=\; {\sim}\mathbb{T}\lambda$	SIL
(5)	$\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner \rightarrow {\sim}\mathbb{T}\lambda \;.{=}.\; {\sim}\mathbb{T}\lambda$	2,4,IL
(6)	$\mathbb{T}\ulcorner\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner \rightarrow {\sim}\mathbb{T}\lambda\urcorner \;=\; \mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner$	5,SIL
(7)	$\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner$	1,6,SIL
(8)	$\mathbb{T}\lambda$	2,7,IL
(9)	${\sim}\mathbb{T}\lambda$	7,direct-**dis**-quotation

Although general alethiation is not admitted by $\mathcal{T}$, a weaker version is.

(dr1) $\{\text{FOL}\vdash\phi\} \hookrightarrow \mathbb{T}\ulcorner\phi\urcorner$

Here, $\{\text{FOL}\vdash\phi\}$ is a previous sequence of lines that form a first-order logic proof of ϕ. This is a derived rule, which is obtained by adding the following axiom.

(a39) $\Box\phi \rightarrow \mathbb{T}\ulcorner\phi\urcorner$

Notice carefully, however, that the following stronger formula is not valid.

✕ $\boxdot\phi \rightarrow \mathbb{T}\ulcorner\phi\urcorner$

26.2. The K-Principle and the D-Principle

The K-principle can be successfully transferred to System $\mathcal{T}$. In particular, we postulate the following *stronger* principle as an axiom, with the desired theorem as an immediate corollary.

(a40) $\mathbb{T}\ulcorner\phi \rightarrow \psi\urcorner \rightarrow \mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\urcorner \rightarrow \mathbb{T}\ulcorner\psi\urcorner\urcorner$

(t83) $\mathbb{T}\ulcorner\phi \rightarrow \psi\urcorner \rightarrow.\; \mathbb{T}\ulcorner\phi\urcorner \rightarrow \mathbb{T}\ulcorner\psi\urcorner$

Next, the alethic counterpart of the D-principle is, as expected, an immediate consequence of the T-principle.

(t84) $\mathbb{T}\ulcorner\phi\urcorner \;\rightarrow\; {\sim}\mathbb{T}\ulcorner{\sim}\phi\urcorner$

Notice, however, that we cannot introduce the alethiated versions of these principles.

✗ $\mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\rightarrow\psi\urcorner \;\rightarrow.\; \mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\urcorner\rightarrow\mathbb{T}\ulcorner\psi\urcorner\urcorner\urcorner$

✗ $\mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\rightarrow\psi\urcorner \;\rightarrow.\; \mathbb{T}\ulcorner\phi\urcorner\rightarrow\mathbb{T}\ulcorner\psi\urcorner\urcorner$

✗ $\mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\urcorner \;\rightarrow\; {\sim}\mathbb{T}\ulcorner{\sim}\phi\urcorner\urcorner$

The first two do not admit a simple counter-example, but must be shown invalid in the proposed semantics (Part 2). The latter does admit a simple counter-example, obtained by substituting ${\sim}\mathbb{T}\lambda$ for ϕ.

(?) $\mathbb{T}\ulcorner\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner \;\rightarrow\; {\sim}\mathbb{T}\ulcorner{\sim}{\sim}\mathbb{T}\lambda\urcorner\urcorner$

(1)	$\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner$	Pr
(2)	$\mathbb{T}\ulcorner\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner \;\rightarrow\; {\sim}\mathbb{T}\ulcorner{\sim}{\sim}\mathbb{T}\lambda\urcorner\urcorner$	Pr
(3)	~~SHOW~~: ✗	15,16,SL
(4)	${\sim}{\sim}\mathbb{T}\lambda = \mathbb{T}\lambda$	SIL
(5)	$\mathbb{T}\ulcorner{\sim}{\sim}\mathbb{T}\lambda\urcorner = \mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner$	4+SIL
(6)	$\mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner = \mathbb{T}\lambda$	(t81)
(7)	$\mathbb{T}\ulcorner{\sim}{\sim}\mathbb{T}\lambda\urcorner = \mathbb{T}\lambda$	5,6,SIL
(8)	${\sim}\mathbb{T}\ulcorner{\sim}{\sim}\mathbb{T}\lambda\urcorner = {\sim}\mathbb{T}\lambda$	7,SIL
(9)	$\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner = \mathbb{T}\lambda$	1,IL
(10)	$\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner \;\rightarrow\; {\sim}\mathbb{T}\ulcorner{\sim}{\sim}\mathbb{T}\lambda\urcorner \;.\!=.\; \mathbb{T}\lambda \;\rightarrow\; {\sim}\mathbb{T}\lambda$	8,9,SIL
(11)	$\mathbb{T}\ulcorner\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner \;\rightarrow\; {\sim}\mathbb{T}\ulcorner{\sim}{\sim}\mathbb{T}\lambda\urcorner\urcorner \;.\!=.\; \mathbb{T}\ulcorner\mathbb{T}\lambda \;\rightarrow\; {\sim}\mathbb{T}\lambda\urcorner$	10,SIL
(12)	$\mathbb{T}\lambda \;\rightarrow\; {\sim}\mathbb{T}\lambda \;.\!=\; {\sim}\mathbb{T}\lambda$	SIL
(13)	$\mathbb{T}\ulcorner\mathbb{T}\lambda \;\rightarrow\; {\sim}\mathbb{T}\lambda\urcorner \;=\; \mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner$	12,SIL
(14)	$\mathbb{T}\ulcorner\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner \;\rightarrow\; {\sim}\mathbb{T}\ulcorner{\sim}{\sim}\mathbb{T}\lambda\urcorner\urcorner \;=\; \mathbb{T}\lambda$	9,11,13,SIL
(15)	$\mathbb{T}\lambda$	2,14,SIL
(16)	${\sim}\mathbb{T}\lambda$	1+(t66)

26.3. The B-Principle, G-Principle, 5-Principle, and L-Principle

We now consider the alethic versions of the B-principle, G-principle [“Geach”], 5-principle, and L-principle [“linear”]. As we see, each one runs afoul of the Liar Paradox, and so cannot be added to System $\mathcal{T}$.

✗ (B) ${\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\phi\urcorner\urcorner \;\rightarrow\; \phi$

✗ (G) ${\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\phi\urcorner\urcorner \;\rightarrow\; \mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner{\sim}\phi\urcorner\urcorner$

✗ (5) ${\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\phi\urcorner\urcorner \;\rightarrow\; \mathbb{T}\ulcorner\phi\urcorner$

✗ (L) $\mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\urcorner \;\rightarrow\; \psi\urcorner \;\vee\; \mathbb{T}\ulcorner\mathbb{T}\ulcorner\psi\urcorner \;\rightarrow\; \phi\urcorner$

In the case of the B-principle, consider substituting ‘$\mathbb{T}\lambda$’ for ϕ, resulting in the following instance.

(?) ${\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner\urcorner \;\rightarrow\; \mathbb{T}\lambda$

Then we can produce a contradiction granted the existence of a liar sentence.

(17)	$\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner$	Pr
(18)	${\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner\urcorner \rightarrow \mathbb{T}\lambda$	Pr
(19)	~~SHOW~~: ✕	9,10,SL
(20)	$\mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner = \mathbb{T}\lambda$	(t81)
(21)	${\sim}\mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner = {\sim}\mathbb{T}\lambda$	4,SIL
(22)	$\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner\urcorner = \mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner$	5,SIL
(23)	$\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner\urcorner = \mathbb{T}\lambda$	1,6,IL
(24)	${\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner\urcorner = {\sim}\mathbb{T}\lambda$	7,SIL
(25)	${\sim}\mathbb{T}\lambda \rightarrow \mathbb{T}\lambda$	2,8,SIL
(26)	${\sim}\mathbb{T}\lambda$	1+(t66)

In the case of the G-principle, also consider substituting '$\mathbb{T}\lambda$' for ϕ, resulting in the following instance.

(?) $\quad {\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner\urcorner \rightarrow \mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner\urcorner$

The reasoning is fairly similar to the case of the B-principle; it shows that the G-principle is inconsistent with the existence of a liar sentence

The alethic version of 5-principle also runs afoul of the Liar Paradox. In particular, consider substituting '${\sim}\mathbb{T}\lambda$' for ϕ; then we have the following instance.

(?) $\quad {\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner\urcorner \rightarrow \mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner$

Once again, we can produce a contradiction granted the existence of a liar sentence.

(1)	$\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner$	Pr
(2)	${\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner\urcorner \rightarrow \mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner$	Pr
(3)	~~SHOW~~: ✕	10,11,SL
(4)	$\mathbb{T}\lambda = \mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner$	1,IL
(5)	${\sim}\mathbb{T}\lambda = {\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner$	4,SIL
(6)	$\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner = \mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner\urcorner$	5,SIL
(7)	$\mathbb{T}\lambda = \mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner\urcorner$	1,6,IL
(8)	${\sim}\mathbb{T}\lambda = {\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner\urcorner$	7,SIL
(9)	${\sim}\mathbb{T}\lambda \rightarrow \mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner$	2,8,SIL
(10)	${\sim}\mathbb{T}\lambda \rightarrow \mathbb{T}\lambda$	1,9,IL
(11)	${\sim}\mathbb{T}\lambda$	1+(t66)

Finally, we note that the L-principle has the following as a special case,

✕ $\quad \mathbb{T}\ulcorner\mathbb{T}\ulcorner\phi\urcorner \rightarrow \phi\urcorner$

which we have already rejected in Section 26.1.

26.4. The Anti-5 Principle, The Anti-B-Principle, and The B-Minus-Principle

We next consider principles that might be called "anti-5" and "anti-B".

(–5) $\quad \Box{\sim}\Box{\sim}\phi \rightarrow {\sim}\Box{\sim}\Box\phi$

Adding this strange principle [which is the converse of the G-principle] to System KT5 (a.k.a. S5) results in complete modal collapse[41] [hence the proposed name 'anti-5']. Similarly adding the following principle

(–B) $\Box{\sim}\phi \rightarrow \Box{\sim}\Box\phi$

to System KTB (a.k.a. B) results in complete modal collapse.

While we are at it, we also consider the following weakening of the B-principle, which we call "B-minus".

(B^{-}) $\Box{\sim}\Box{\sim}\Box\phi \rightarrow \Box\phi$

As it turns out, the alethic counterparts of *all* these modal principles are valid in $\mathcal{T}$. Indeed, adding the following axiom yields all of them.

(a41) $\langle \mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\phi\urcorner\urcorner\rangle = \langle \mathbb{T}\ulcorner{\sim}\phi\urcorner\rangle$

We then have the following as theorems, as well as their alethiations.

(t85) $\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner{\sim}\phi\urcorner\urcorner \rightarrow {\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\phi\urcorner\urcorner$

(t86) $\mathbb{T}\ulcorner{\sim}\phi\urcorner \rightarrow \mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\phi\urcorner\urcorner$

(t87) $\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\ulcorner\phi\urcorner\urcorner\urcorner \rightarrow \mathbb{T}\phi$

26.5. The Modal-Collapse Principle

The next item from the above list is C – the principle of modal "collapse". Adding it to *most* modal systems results in a trivial modal system (i.e., modal collapse). The alethic counterpart is rejected by $\mathcal{T}$.

✕ ${\sim}\mathbb{T}\ulcorner{\sim}\phi\urcorner \rightarrow \mathbb{T}\ulcorner\phi\urcorner$

It's easy to see that we cannot add this principle to $\mathcal{T}$. For consider substituting $\mathbb{T}\lambda$ for ϕ, where $\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner$.

(1)	${\sim}\mathbb{T}\ulcorner{\sim}\mathbb{T}\lambda\urcorner \rightarrow \mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner$	Pr
(2)	$\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner$	Pr
(3)	~~SHOW~~: ✕	4,7,SL
(4)	${\sim}\mathbb{T}\lambda$	2,et
(5)	${\sim}\mathbb{T}\lambda \rightarrow \mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner$	1,2,IL
(6)	$\mathbb{T}\ulcorner\mathbb{T}\lambda\urcorner$	4,5,SL
(7)	$\mathbb{T}\lambda$	6,direct-disquotation

Notice also that the above principle is immediately equivalent to the following.

[41] Modal collapse occurs when a simple modal operator is logically redundant; e.g., $\Box\phi = \phi$, or $\Diamond\phi = \phi$, or $\Box\phi = \Diamond\phi$. Complete modal collapse occurs when all of these "identities" obtain.

(?) $\mathbb{T}\ulcorner\phi\urcorner$ or $\mathbb{T}\ulcorner{\sim}\phi\urcorner$

i.e., every sentence or its negation is true.

We examine this in great detail in Section 27 (“Truth-Value Gaps”).

26.6. The Barcan Formula

The final item on our list is bf – the Barcan Formula. What about its alethic counterpart.

✕ $\mathbb{T}\ulcorner\forall\nu\phi\urcorner \leftrightarrow \forall\nu\mathbb{T}\ulcorner\phi\urcorner$

This cannot be added as a principle of alethic logic. The chief reason it is not valid is that, on the right side, we are *attempting* to referentially quantify into direct-quotes, but direct-quotes are referentially opaque.[42]

We can, however, salvage half of the Barcan formula, via the following theorem schema.

(t88) $\mathbb{T}\ulcorner\forall\nu\phi\urcorner \rightarrow \mathbb{T}\ulcorner\phi[\boldsymbol{t}/\nu]\urcorner$

Here, $\phi[\boldsymbol{t}/\nu]$ is the formula that results from substituting the singular term $\boldsymbol{t}$ for every occurrence of ν in ϕ that is free for $\boldsymbol{t}$.

27. Truth-Value Gaps

27.1. Definitions

A *truth-value gap* is usually defined to be a semantic situation in which a sentence/proposition has *no truth-value*; it is neither true nor false. The natural location of this notion is in the meta-theory. Nevertheless, since the object theory $\mathcal{T}$ has a truth-predicate, we can also implement the notion in the object language. We adopt the term ‘gapped’ [or ‘gappy’] — symbolized by ‘$\mathbb{G}$’ in the object language.

The following seems to be a natural definition.

(d21) $\mathbb{G}\boldsymbol{x} =_{\text{df}} \mathbb{S}\boldsymbol{x}$ or $\mathbb{P}\text{x}$.& $\sim\mathbb{T}\boldsymbol{x}$ & $\sim\mathbb{F}\boldsymbol{x}$

In other words,

A sentence/proposition is “gapped” if and only if it is neither true nor false.

This depends upon a prior definition of ‘…is false’ ($\mathbb{F}$), which we have not yet provided. The following is one plausible definition.

(?) $\mathbb{F}\boldsymbol{x} =_{\text{df}} \mathbb{S}\boldsymbol{x}$ or $\mathbb{P}\boldsymbol{x}$.& $\sim\mathbb{T}\boldsymbol{x}$

[42] Notice, of course, that the indirect-quotation counterpart is perfectly ok.
$\mathbb{T}\langle\forall\nu\phi\rangle \leftrightarrow \forall\nu\mathbb{T}\langle\phi\rangle$

The problem with this definition is that it renders the predicate 'is gapped' completely trivial. It is accordingly preferable to define falsity as follows.

(d22) $\mathbb{F}\boldsymbol{x} =_{\text{df}} \mathbb{S}\boldsymbol{x}$ or $\mathbb{P}\boldsymbol{x}$.& $\mathbb{T}\neg\boldsymbol{x}$

In other words, a sentence/proposition is false if and only if its negation is true.

Therefore, since $\mathbb{F}\boldsymbol{x}$ means $\mathbb{T}\neg\boldsymbol{x}$, our definition of 'gapped' comes to the following.

A sentence/proposition is *gapped* (or *gappy*) if and only if neither it nor its negation is true.

The following are immediate theorems.

(t89) $\mathbb{F}\langle\phi\rangle \leftrightarrow \mathbb{T}\langle\sim\phi\rangle$

(t90) $\mathbb{G}\langle\phi\rangle \leftrightarrow.\ \sim\mathbb{T}\langle\phi\rangle\ \&\ \sim\mathbb{T}\langle\sim\phi\rangle$

(t91) $\mathbb{F}\ulcorner\phi\urcorner \leftrightarrow \mathbb{T}\ulcorner\sim\phi\urcorner$

(t92) $\mathbb{G}\ulcorner\phi\urcorner \leftrightarrow.\ \sim\mathbb{T}\ulcorner\phi\urcorner\ \&\ \sim\mathbb{T}\ulcorner\sim\phi\urcorner$

27.2. There are no Gappy Propositions

Next, we observe the following key theorem and corollaries.

(t93) $\mathbb{T}\langle\phi\rangle$ or $\mathbb{T}\langle\sim\phi\rangle$

(t94) $\forall p\{\mathbb{T}p$ or $\mathbb{T}\neg p\}$

(t95) $\sim\exists p\mathbb{G}p$

In other words,

Every proposition or its negation is true; there are no gappy propositions.

Notice that (t93) is an immediate consequence of the indirect-**en**-quotation principle.

(1)	$\phi \rightarrow \mathbb{T}\langle\phi\rangle$	indirect-**en**-quotation
(2)	$\sim\phi \rightarrow \mathbb{T}\langle\sim\phi\rangle$	indirect-**en**-quotation
(3)	$\mathbb{T}\langle\phi\rangle$ or $\mathbb{T}\langle\sim\phi\rangle$	1,2,SL

27.3. There are Gappy Sentences

Given theorem (t93), which concerns indirect-quotation, one might expect the corresponding principle concerning direct-quotation to be valid. In other words, one might expect the following to be theorems.

✕ $\mathbb{T}\ulcorner\phi\urcorner$ or $\mathbb{T}\ulcorner{\sim}\phi\urcorner$

✕ $\forall s\{\mathbb{T}s$ or $\mathbb{T}\neg s\}$

✕ ${\sim}\exists s\mathbb{G}s$

Unfortunately, this is not so, as explained in the following meta-theorem.

(mt1) $\forall\phi\vdash\{\mathbb{T}\ulcorner\phi\urcorner$ or $\mathbb{T}\ulcorner{\sim}\phi\urcorner\} \rightarrow \forall\phi\vdash\{\phi\rightarrow\mathbb{T}\ulcorner\phi\urcorner\}$

In other words:

[Granting direct-**dis**-quotation,]
we cannot rid ourselves of truth-value gaps without accepting direct-**en**-quotation.
If we abandon direct-**en**-quotation, then we must admit truth-value gaps.

The following simple proof shows why.

(1)	$\mathbb{T}\ulcorner\phi\urcorner$ or $\mathbb{T}\ulcorner{\sim}\phi\urcorner$	Pr
(2)	~~SHOW~~: $\phi \rightarrow \mathbb{T}\ulcorner\phi\urcorner$	CD
(3)	| ϕ	As
(4)	| ~~SHOW~~: $\mathbb{T}\ulcorner\phi\urcorner$	As
(5)	| | ${\sim}\mathbb{T}\ulcorner\phi\urcorner$	As
(6)	| | ~~SHOW~~: ✕	3,8,SL
(7)	| | | $\mathbb{T}\ulcorner{\sim}\phi\urcorner$	1,5,SL
(8)	| | | ${\sim}\phi$	7,direct-**dis**-quotation

27.4. The Liar Sentence is Gappy

An immediate example of a gappy sentence is the liar sentence λ, as seen in the following theorem.

(t96) $\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner \rightarrow \mathbb{G}[\lambda]$

(1)	$\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner$	As
(2)	$\neg\lambda = \ulcorner{\sim}{\sim}\mathbb{T}\lambda\urcorner$	1,Def(¬)
(3)	~~SHOW~~: $\mathbb{G}[\lambda]$	3,Def(𝔾)
(4)	~~SHOW~~: $\sim\mathbb{T}\lambda \;\&\; \sim\mathbb{T}\neg\lambda$	4,5,SL
(5)	│ $\sim\mathbb{T}\lambda$	1+(t66)
(6)	│ ~~SHOW~~: $\sim\mathbb{T}\neg\lambda$	ID
(7)	│ │ $\mathbb{T}\neg\lambda$	As
(8)	│ │ ~~SHOW~~: ✕	5,10,SL
(9)	│ │ │ $\mathbb{T}\ulcorner{\sim}{\sim}\mathbb{T}\lambda\urcorner$	2,7,IL
(10)	│ │ │ $\sim\sim\mathbb{T}\lambda$	9,direct-**dis**-quotation

28. The Normal Core of System $\mathcal{T}$

We have already discussed the Liar Paradox and the Deferred Liar Paradox. There are numerous other liar-like paradoxes. However, these paradoxes are better understood after we have a fairly clear idea about what will be called the "normal core" of System $\mathcal{T}$.

First, the basic idea. Although we have rejected the *general* validity of the naive truth-schema

(nts) $\mathbb{T}\ulcorner\phi\urcorner \leftrightarrow \phi$

we nevertheless want to preserve as many of its *useful* instances as we can, including the following.

⌜snow is white⌝ is true ↔ snow *is* white
⌜grass is green⌝ is true ↔ grass *is* green
⌜2+2 is 4⌝ is true ↔ 2+2 *is* 4
etc.

The question is how to accomplish this (while maintaining the consistency of our theory). This is where the notion of *normalcy* is introduced. Our goal is to obtain the following as a theorem of $\mathcal{T}$.

$$\text{normal}[\phi] \rightarrow. \;\mathbb{T}\ulcorner\phi\urcorner \leftrightarrow \phi$$

In other words:

Normal sentences have normal truth conditions.

This is all well and good so long as we provide a characterization of 'normal'.

To start with, we list the following (fairly informal) intuitions about *normalcy*.

(1) most expressions of ordinary English are normal;
(2) propositions exclusively involving ordinary objects are normal;
(3) grounded[43] propositions are normal;
(4) necessary propositions are normal
(5) impossible propositions are normal
(6) identities are normal
(7) ordinary objects are normal;
(8) normal singular terms denote ordinary objects;
(9) normal function signs denote normal functions;
(10) normal predicates denote normal propositional functions;
(11) normal formulas denote normal propositions;
(12) whether a singular term is normal may be a contingent matter;
(13) atomic sentence-denoting terms *need not be* normal;
(14) atomic proposition-denoting terms *need not be* normal;
(15) truth-ascriptions *may or may not be* normal.

In order to implement this notion in our theory, we introduce two new symbols – '$\mathbb{N}$' and $\mathcal{N}$'. Whereas '$\mathbb{N}$' is a primitive one-place predicate, '$\mathcal{N}$' is a primitive multi-variable-binding operator. In the case of the latter, the following schema is postulated.

(a42) $\mathcal{N}\underline{x}\phi \;\rightarrow.\; \mathbb{T}\ulcorner\phi[\underline{t}/\underline{x}]\urcorner = \phi[\underline{t}/\underline{x}]$

Here, '$\underline{x}$' and '$\underline{t}$' are special meta-variables ranging over sequences of variables and singular terms respectively. Note that this includes the empty sequence, so the following is a special case, from which the desired theorem as an immediate consequence.

(t97) $\mathcal{N}\phi \;\rightarrow.\; \mathbb{T}\ulcorner\phi\urcorner = \phi$

(t98) $\mathcal{N}\phi \;\rightarrow.\; \mathbb{T}\ulcorner\phi\urcorner \leftrightarrow \phi$

Furthermore, the following are postulated by $\mathcal{T}$.

(a43) $\mathbb{N}[\boldsymbol{x}] \rightarrow \exists y[\boldsymbol{x} = y]$

(a44) $\forall\underline{x}\Box\phi \rightarrow \mathcal{N}\underline{x}\phi$

(a45) $\forall\underline{x}\Box{\sim}\phi \rightarrow \mathcal{N}\underline{x}\phi$

(a46) $\mathbb{O}[x] \rightarrow \mathbb{N}[x]$

(a47) $\forall\underline{x}[\phi{=}\psi] \;.\!\rightarrow.\; \mathcal{N}\underline{x}\phi = \mathcal{N}\underline{x}\psi$

(a48) $\mathcal{N}\underline{x}\phi \;\&\; \mathcal{N}\underline{x}\psi \;.\!\rightarrow\; \mathcal{N}\underline{x}(\phi \;\&\; \psi)$

[43] Here, of course, I have in mind Kripke's notion of *grounded*. However, *grounded* and *normal* are not coextensive. We presume that many un-grounded sentences are nevertheless normal and therefore have normal truth-conditions.

(a49) $\forall\nu\mathcal{N}\underline{x}\phi \rightarrow \mathcal{N}\underline{x}\forall\nu\phi$

(a50) $\mathcal{N}\phi$ [provided ϕ contains no constant, proper noun, or free variable]

(a51) $\mathbb{N}[\ulcorner\phi\urcorner] \leftrightarrow \mathcal{N}\phi$

(a52) $\mathbb{N}[\langle\phi\rangle] \leftrightarrow \mathcal{N}\phi$

(a53) $\mathbb{N}[x_1]\ \&\ \ldots\ \&\ \mathbb{N}[x_k]\ .\!\rightarrow\ \mathcal{N}\boldsymbol{P}[x_1,\ldots,x_k]\ \&\ \mathcal{N}{\sim}\boldsymbol{P}[x_1,\ldots,x_k]$

For example, the following are *plausible* theorems of $\mathcal{T}$, granting *plausible* assumptions stated to the right.

(pt1) $\mathcal{N}$\{snow is white\} [‘snow’ is ordinary]

(c) ⌜snow is white⌝ is true ↔ snow is white

(pt2) $\mathcal{N}$\{snow is true\} [‘snow’ is ordinary]

(c) ⌜snow is true⌝ is true ↔ snow is true

(pt3) $\mathcal{N}$\{grass is green\} [‘grass’ is ordinary]

(c) ⌜grass is green⌝ is true ↔ grass is green

Whereas the above items, and countless others like them, are *highly plausible* theorems, the following are genuine theorems.

(t99) $\mathcal{N}\{2+2=4\}$

(c) $\ulcorner 2+2=4\urcorner$ is true $\leftrightarrow$ $2+2=4$

(t100) $\mathcal{N}\{\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner\}$

(c) $\ulcorner\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner\urcorner$ is true $\leftrightarrow$ $\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner$

(t101) $\mathcal{N}\{\exists s{\sim}\mathbb{T}s\}$

(c) $\ulcorner\exists s{\sim}\mathbb{T}s\urcorner$ is true $\leftrightarrow$ $\exists s{\sim}\mathbb{T}s$

On the other hand, whether we have the following

(?) $\mathcal{N}({\sim}\mathbb{T}\lambda)$

will depend upon what ‘λ’ denotes. In particular, we have the following.

(t102) $\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner \rightarrow {\sim}\mathbb{N}[\lambda]$

(t103) $\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner \rightarrow {\sim}\mathcal{N}({\sim}\mathbb{T}\lambda)$

The liar sentence is *not* normal!

29. Contingently Paradoxical Sentences

Having discussed the normal core of $\mathcal{T}$, we now return to the discussion of the variety of paradoxes associated with the concept of truth.

Kripke was probably the first to point out that whether a statement is paradoxical may depend upon contingent empirical matters. This means, he argues, that we cannot develop a simple-minded syntactic or semantic procedure to winnow out sentences that will cause serious trouble for the naive truth-schema.

We will not specifically discuss Kripke's examples, but will instead look at two simple single-sentence examples that illustrate how a sentence can be contingently paradoxical. The first one is given as follows.

(ρ) it is raining, and this sentence is not true

Here, it is presumed that 'it is raining' is normal (and contingent).

First, we formalize the situation as follows.

($\rho^{=}$) $s = \ulcorner R \,\&\, \sim\mathbb{T}s \urcorner$

Next, we show that, according to the naive truth-schema, we have a contradiction whenever it is raining.

(1)	$s = \ulcorner R \,\&\, \sim\mathbb{T}s \urcorner$	Pr
(2)	R	Pr
(3)	~~SHOW~~: ✕	4,11,SL
(4)	│ ~~SHOW~~: $\mathbb{T}s$	ID
(5)	│ │ $\sim\mathbb{T}s$	As
(6)	│ │ ~~SHOW~~: ✕	5,9,SL
(7)	│ │ │ $R \,\&\, \sim\mathbb{T}s$	2,5,SL
(8)	│ │ │ $\mathbb{T}\ulcorner R \,\&\, \sim\mathbb{T}s \urcorner$	7, ✕direct-en-quotation✕
(9)	│ │ │ $\mathbb{T}s$	1,8,IL
(10)	│ $\mathbb{T}\ulcorner R \,\&\, \sim\mathbb{T}s \urcorner$	1,4,IL
(11)	│ $R \,\&\, \sim\mathbb{T}s$	10, direct-dis-quotation

We next observe that, if we assume instead that it is not raining ($\sim R$), then no contradiction can be obtained. Therefore, we conclude that, depending on whether it is or is not raining, we have a paradox (i.e., contradiction).

System $\mathcal{T}$ describes the above situation in very similar language. In particular, we will say that a semantic situation is paradoxical when a sentence is gapped. Note carefully that this not a paradox in the sense of a contradiction; it is merely a puzzling situation. Let us demonstrate this by showing how $\mathcal{T}$ deals with the above example.

(1)	$s = \ulcorner R \,\&\, {\sim}\mathbb{T}s \urcorner$	Pr
(2)	R	Pr
(3)	~~SHOW~~: $\mathbb{G}s$	ID
(4)	$\quad {\sim}\mathbb{G}s$	As
(5)	$\quad$ ~~SHOW~~: ✕	sep cases
(6)	$\quad\quad \mathbb{T}s$ or $\mathbb{T}\neg s$	4,Def($\mathbb{G}$)
(7)	$\quad$ c1: $\mathbb{T}s$	As
(8)	$\quad$ ~~SHOW~~: ✕	7,10,SL
(9)	$\quad\quad \mathbb{T}\ulcorner R \,\&\, {\sim}\mathbb{T}s \urcorner$	1,9,IL
(10)	$\quad\quad R \,\&\, {\sim}\mathbb{T}s$	9,direct-**dis**-quotation
(11)	$\quad$ c2: $\mathbb{T}\neg s$	As
(12)	$\quad$ ~~SHOW~~: ✕	17,19,SL
(13)	$\quad\quad \neg s = \neg \ulcorner R \,\&\, {\sim}\mathbb{T}s \urcorner$	1,IL
(14)	$\quad\quad \neg \ulcorner R \,\&\, {\sim}\mathbb{T}s \urcorner = \ulcorner {\sim}(R \,\&\, {\sim}\mathbb{T}s) \urcorner$	ET
(15)	$\quad\quad \mathbb{T}\ulcorner {\sim}(R \,\&\, {\sim}\mathbb{T}s) \urcorner$	11,13,14,IL
(16)	$\quad\quad {\sim}(R \,\&\, {\sim}\mathbb{T}s)$	15,direct-**dis**-quotation
(17)	$\quad\quad \mathbb{T}s$	2,16,SL
(18)	$\quad\quad \mathbb{T}\ulcorner R \,\&\, {\sim}\mathbb{T}s \urcorner$	1,17,IL
(19)	$\quad\quad R \,\&\, {\sim}\mathbb{T}s$	18,direct-**dis**-quotation

This tells us that, if it is raining, then the sentence ‘it is raining and this sentence is not true’ is gapped (paradoxical, if you like). What about if it is *not* raining? Since we suppose that ‘it is raining’ is normal, from the assumption that it is raining we can deduce that the sentence ‘it is not raining’ is true. Watch for this step in the following proof.

(1)	$s = \ulcorner R \,\&\, {\sim}\mathbb{T}s \urcorner$	Pr
(2)	${\sim}R$	Pr
(3)	~~SHOW~~: ${\sim}\mathbb{G}s$	4,Def($\mathbb{G}$)
(4)	~~SHOW~~: $\mathbb{T}\neg s$	5,6,IL
(5)	$\neg s = \ulcorner {\sim}(R \,\&\, {\sim}\mathbb{T}s) \urcorner$	1,Def($\neg$)
(6)	~~SHOW~~: $\mathbb{T}\ulcorner {\sim}(R \,\&\, {\sim}\mathbb{T}s) \urcorner$	2,8,$\leqslant$-OUT
(7)	$\quad \mathbb{T}\ulcorner {\sim}R \urcorner$	2+normality of R
(8)	$\quad {\sim}R \leqslant {\sim}(R \,\&\, {\sim}\mathbb{T}s)$	SL result??
(9)	$\quad \mathbb{T}\ulcorner {\sim}\mathbb{R} \urcorner \leqslant \mathbb{T}\ulcorner {\sim}(R \,\&\, {\sim}\mathbb{T}s) \urcorner$	7+monotonicity

We can combine and generalize our result in the following theorem schema.

(t104) $s = \ulcorner \phi \,\&\, {\sim}\mathbb{T}s \urcorner \;\rightarrow.\; \mathbb{G}s \leftrightarrow \phi$ [one instance for each normal formula ϕ]

30. Löb’s Paradox

In this section, we examine Löb’s Paradox. First, we will use it to “prove” that 2+2=5! Consider the sentence

(γ) if this sentence is true, then 2+2=5.

As usual, we first formalize this sentence, as follows.

($\gamma^{=}$) $s = \ulcorner \mathbb{T}s \rightarrow 2{+}2{=}5 \urcorner$

Given this premise, we can produce some more mischief.

(1)	$s = \ulcorner \mathbb{T}s \rightarrow 2{+}2{=}5 \urcorner$	Pr
(2)	~~SHOW~~: $2{+}2{=}5$	DD
(3)	~~SHOW~~: $\mathbb{T}s \rightarrow 2{+}2{=}5$	CD
(4)	$\mathbb{T}s$	As
(5)	~~SHOW~~: $2{+}2{=}5$	DD
(6)	$\mathbb{T}\ulcorner \mathbb{T}s \rightarrow 2{+}2{=}5 \urcorner$	1,4,IL
(7)	$\mathbb{T}s \rightarrow 2{+}2{=}5$	6,**dis**-quotation
(8)	$2{+}2{=}5$	4,7,SL
(9)	$\mathbb{T}\ulcorner \mathbb{T}s \rightarrow 2{+}2{=}5 \urcorner$	3, ✕**en**-quotation✕
(10)	$\mathbb{T}s$	1,9,IL
(11)	$2{+}2{=}5$	3,10,SL

Thus, we have shown that 2+2=5 based on the mere *existence* of sentence γ! Of course, as marked by '✕' in the annotation column, there is an invalid step, which appeals to direct **en**-quotation.

There is a valid core to this proof, however. In particular, we can prove the following schema.

(t105) $s = \ulcorner \mathbb{T}s \rightarrow \phi \urcorner \;\rightarrow.\; \mathbb{T}s \rightarrow \phi$

(1)	$s = \ulcorner \mathbb{T}s \rightarrow \phi \urcorner$	Pr
(2)	~~SHOW~~: $\mathbb{T}s \rightarrow \phi$	CD
(3)	$\mathbb{T}s$	As
(4)	~~SHOW~~: ϕ	DD
(5)	$\mathbb{T}\ulcorner \mathbb{T}s \rightarrow \phi \urcorner$	1,3,IL
(6)	$\mathbb{T}s \rightarrow \phi$	5,**dis**-quotation
(7)	ϕ	3,6,SL

Like our previous sentence ρ, the sentence γ may be viewed as contingently paradoxical. In particular, if ϕ is true, then according to NTT, γ is unproblematically true. But if ϕ is not true, then according to NTT, the existence of γ leads to paradox (i.e., contradiction).

Our own analysis is terminologically similar. In particular, we have the following theorem schema — that the Löb sentence is paradoxical (i.e., gapped) exactly when $\ulcorner\phi\urcorner$ is not true.

(t106) $s = \ulcorner \mathbb{T}s \rightarrow \phi \urcorner \;\rightarrow.\; \mathbb{G}[s] \leftrightarrow {\sim}\mathbb{T}\ulcorner\phi\urcorner$

(1)	$s = \ulcorner \mathbb{T}s \rightarrow \phi \urcorner$	As
(2)	$\neg s = \neg\ulcorner \mathbb{T}s \rightarrow \phi \urcorner = \ulcorner {\sim}(\mathbb{T}s \rightarrow \phi) \urcorner$	1,IL/Def($\neg$)
(3)	~~SHOW~~: $\mathbb{G}[s] \leftrightarrow {\sim}\mathbb{T}\ulcorner\phi\urcorner$	
	[$\rightarrow$]	
(4)	$\mathbb{G}[s]$	As
(5)	~~SHOW~~: ${\sim}\mathbb{T}\ulcorner\phi\urcorner$	ID
(6)	$\mathbb{T}\ulcorner\phi\urcorner$	As
(7)	~~SHOW~~: ✕	12,13,SL
(8)	$\phi \rightarrow . \mathbb{T}s \rightarrow \phi$	SL($\vdash$)
(9)	$\ulcorner\phi\urcorner \leqslant \ulcorner \mathbb{T}s \rightarrow \phi \urcorner$	8+??
(10)	$\mathbb{T}\ulcorner\phi\urcorner \leqslant \mathbb{T}\ulcorner \mathbb{T}s \rightarrow \phi \urcorner$	9,mono
(11)	$\mathbb{T}\ulcorner \mathbb{T}s \rightarrow \phi \urcorner$	6,10,$\leqslant$-OUT
(12)	$\mathbb{T}s$	1,11,IL
(13)	${\sim}\mathbb{T}s \;\&\; {\sim}\mathbb{F}s$	4,Def($\mathbb{G}$)
	[$\leftarrow$]	

(14)	$\sim\mathbb{T}\ulcorner\phi\urcorner$	As
(15)	~~SHOW~~: $\mathbb{G}[s]$	16,Def($\mathbb{G}$)
(16)	~~SHOW~~: $\sim\mathbb{T}s \;\&\; \sim\mathbb{T}[\neg s]$	17,22,SL
(17)	~~SHOW~~: $\sim\mathbb{T}s$	ID
(18)	$\mid\; \mathbb{T}s$	As
(19)	$\mid\;$ ~~SHOW~~: ✕	14,25,SL
(20)	$\mid\;\mid\; \mathbb{T}s = \mathbb{T}\ulcorner\mathbb{T}s\rightarrow\phi\urcorner$	1,IL
(21)	$\mid\;\mid\; \mathbb{T}\ulcorner\mathbb{T}s\rightarrow\phi\urcorner \precsim \mathbb{T}\ulcorner\mathbb{T}s\urcorner\rightarrow\mathbb{T}\ulcorner\phi\urcorner$	??
(22)	$\mid\;\mid\; \mathbb{T}s \precsim \mathbb{T}\ulcorner\mathbb{T}s\urcorner\rightarrow\mathbb{T}\ulcorner\phi\urcorner$	20,21,SIL
(23)	$\mid\;\mid\; \mathbb{T}s \precsim \mathbb{T}\ulcorner\mathbb{T}s\urcorner$	??
(24)	$\mid\;\mid\; \mathbb{T}s \precsim \mathbb{T}\ulcorner\phi\urcorner$	22,23+??
(25)	$\mid\;\mid\; \mathbb{T}\ulcorner\phi\urcorner$	18,24,$\precsim$-OUT
(26)	~~SHOW~~: $\sim\mathbb{T}[\neg s]$	ID
(27)	$\mid\; \mathbb{T}[\neg s]$	As
(28)	$\mid\;$ ~~SHOW~~: ✕	17,31,SL
(29)	$\mid\;\mid\; \mathbb{T}\ulcorner\sim(\mathbb{T}s\rightarrow\phi)\urcorner$	2,27,IL
(30)	$\mid\;\mid\; \sim(\mathbb{T}s\rightarrow\phi)$	29,direct-**dis**-quotation
(31)	$\mid\;\mid\; \mathbb{T}s$	30,SL

With this theorem in hand, we can prove the following immediate corollary.

(t107) $s = \ulcorner\mathbb{T}s \rightarrow \phi\urcorner \;\&\; \mathcal{N}\phi \;\;.\!\rightarrow\!.\; \mathbb{G}[s] \leftrightarrow \sim\phi$

The following theorems amplify some of the ideas in the previous two results.

(t108) $s = \ulcorner\mathbb{T}s \rightarrow \phi\urcorner \;\rightarrow\; \mathbb{T}s \approx \mathbb{T}\ulcorner\phi\urcorner$

(t109) $s = \ulcorner\mathbb{T}s \rightarrow \phi\urcorner \;\rightarrow\; \omega(s) \approx \omega(\ulcorner\phi\urcorner)$

(t110) $\mathcal{N}\phi \;\rightarrow\!.\; s = \ulcorner\mathbb{T}s \rightarrow \phi\urcorner \;\rightarrow\; \mathbb{T}s \approx \phi$

(t111) $\mathcal{N}\phi \;\rightarrow\!.\; s = \ulcorner\mathbb{T}s \rightarrow \phi\urcorner \;\rightarrow\; \omega(s) \approx \langle\phi\rangle$

As we saw earlier (Section 28), most formulas are normal; abnormality is the rare exception. Among the normal formulas is the formula [2+2=5]. So, a goofy sentence like

(γ) if this sentence is true, then 2+2=5

is equivalent to:

$2+2 = 5$

But since the latter is necessarily not true, the original sentence γ is also necessarily not true.

31. The Liar Paradox is a Special Case of Löb’s Paradox

Before continuing, we note that the Liar Paradox is a *special case* of Löb’s Paradox. The basic argument proceeds as follows. First, the following is a theorem of first-order logic.

$$\sim\mathbb{T}s \;\leftrightarrow\!.\; \mathbb{T}s \rightarrow [x\neq x]$$

Accordingly, the following is a theorem of $\mathcal{T}$.

(t112) $\sim\mathbb{T}s \ =. \ \mathbb{T}s \rightarrow [x{\neq}x]$

This allows us to prove the following theorem.

(t113) $\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner \ \rightarrow. \ \mathbb{T}\lambda = \mathbb{T}\ulcorner\mathbb{T}\lambda \rightarrow x{\neq}x\urcorner$

Noting that [x≠x] is impossible, and hence normal, we can apply (t110) to obtain the following.

(t114) $\lambda = \ulcorner{\sim}\mathbb{T}\lambda\urcorner \ \rightarrow. \ \mathbb{T}\lambda \approx [x{\neq}x]$

So, once again, we see how the liar sentence is not true.

32. Pitfalls in the Logic of Truth [+++ to be completed +++]

32.1. The Denotation of a Sentence [+++ to be completed +++]

✕ $\delta(\ulcorner\phi\urcorner) = \langle\phi\rangle$

✕ $\forall s\exists p[p = \delta(s)]$

The problem is the sentence

$$\lambda = \ulcorner{\sim}\mathbb{T}[\delta(\lambda)]\urcorner$$

32.2. What a Sentence Says [+++ to be completed +++]

✕ what($\ulcorner\phi\urcorner$ says) $=$ that(ϕ & $\ulcorner\phi\urcorner$ is true)

The problem is the sentence

$$\lambda = \ulcorner{\sim}\mathbb{T}[\omega(\lambda)]\urcorner$$

PART 2 – SEMANTICS FOR SYSTEM $\mathcal{T}$

33. Introduction

Although the formal theory $\mathcal{T}$ contains its own truth-predicate, we still need to provide a meta-theoretical semantics for $\mathcal{T}$, if for no other reason than to demonstrate that $\mathcal{T}$ and crucial extensions of $\mathcal{T}$ are consistent.[44]

34. The Universe of Discourse

In constructing models of $\mathcal{T}$ (and $\mathcal{T}$-languages), the universe (domain) of discourse, $\mathbb{U}$, will include both "sentences" and "propositions" as well as "ordinary objects", which are all "things" other than sentences and propositions.

Since any $\mathcal{T}$-language $\mathcal{L}$ presumably talks about itself, the intended extension of the predicate 'is a sentence' (the domain of sentences) consists precisely of all the formulas of $\mathcal{L}$ [this set is denoted $\mathbb{S}$]. In this matter, we have no choice, it seems. Concerning propositions, however, there is more room to maneuver. In the proposed semantics, propositions are given *abstractly* by a (complete) π-algebra $\mathbb{P}$ (see Part 3). It is furthermore postulated that no element of $\mathbb{P}$ is an element of $\mathbb{S}$. That leaves "ordinary objects". Here, the semantics simply proposes that the domain $\mathbb{O}$ of ordinary objects is just the set of all remaining entities. In other words, $\mathbb{O} = \mathbb{U}–(\mathbb{S}\cup\mathbb{P})$. From all this it follows, of course, that $\mathbb{U} = \mathbb{S}\cup\mathbb{P}\cup\mathbb{O}$.

Summary

(1)	the domain of sentences	=	$\mathbb{S}$	[the formulas of $\mathcal{L}$]
(2)	the domain of propositions	=	$\mathbb{P}$	[a complete π-algebra]
(3)	$\mathbb{S} \perp \mathbb{P}$			[no sentence is a proposition]
(4)	the domain of ordinary objects	=	$\mathbb{U} – (\mathbb{S}\cup\mathbb{P})$	
(5)	the universe of discourse	=	$\mathbb{U}$	[= $\mathbb{S}\cup\mathbb{P}\cup\mathbb{O}$]

35. The Basic Semantic Apparatus

A formal semantics for a formal language provides a mathematical definition of a class of admissible semantic-valuation functions (or simply valuations). For the sake of simplifying our terminology, we propose to say that

> expression ε *denotes* object o (according to valuation function υ)
> precisely when $\upsilon(\varepsilon)=o$.

[44] In particular, $\mathcal{T}$ + {s = 's is not true'}.

In conventional semantics, every singular term of the formal language $\mathcal{L}$ denotes an element of $\mathbb{U}$, and every formula denotes a truth-value. In the proposed semantics, although we preserve the conventional manner of interpreting singular terms, we generalize the manner of interpreting formulas. In particular, according to the proposed semantics, a formula denotes a proposition (element of $\mathbb{P}$), which in turn is either true or false.[45] This choice of model-object for interpreting formulas influences all the remaining choices, which we summarize as follows.

1. Every formula denotes a proposition (element of $\mathbb{P}$).
2. The elements of $\mathbb{P}$ are divided into the true (T) and the false (F); T forms a maximal complete π-filter on $\mathbb{P}$; F is it complement. [See Part 3.]
3. Among the elements of $\mathbb{P}$ are a special sub-class $\mathcal{N}$ of normal propositions, which form a complete Boolean sub-algebra of $\mathbb{P}$. [See Part 3.]
4. A semantic valuation υ *satisfies* a formula ϕ [i.e., $\upsilon \models \phi$] iff $\upsilon(\phi) \in T$.
5. Every singular term denotes an element of $\mathbb{U}$[46].
   - 5.1. every sentence-denoting term denotes an element of $\mathbb{S}$;
   - 5.2. every proposition-denoting term denotes an element of $\mathbb{P}$;
   - 5.3. every ordinary-denoting term denotes an element of $\mathbb{U}-(\mathbb{S}\cup\mathbb{P})$;
6. Every n-place function sign denotes an n-place function[47] on $\mathbb{U}$.
7. Every n-place predicate P denotes an n-place function[48] – P* – from $\mathbb{U}$ into $\mathbb{P}$, subject to the following restrictions.
   - 7.1. if $a_1,\ldots,a_n \in \mathbb{O}$, then $P^*(a_1,\ldots,a_n) \in \mathcal{N}$
   - 7.2. if P is normal, then $P^*(a_1,\ldots,a_n) \in \mathcal{N}$ for every $a_1,\ldots,a_n$ in $\mathbb{U}$
   - 7.3. $glb(\mathbb{T} \cap \{P^*(x) : x \in \mathbb{U}\}) \in \mathcal{N}$ [*glb* is the greatest lower bound in $\mathbb{P}$]
   - 7.4. $lub(\mathbb{T} \cap \{P^*(x) : x \in \mathbb{U}\}) \in \mathcal{N}$ [*lub* is the least upper bound in $\mathbb{P}$]
8. Every SL connective denotes its corresponding Boolean-function on $\mathbb{P}$.
9. Except for direct-quote-expressions[49] and quantified formulas[50], which get special treatment, the denotation of a compound expression is compositionally (algebraically) constructed from the denotations of its immediate parts. Examples:
   - 9.1. $\upsilon(\sim\alpha) \quad = \quad \upsilon(\sim)[\upsilon(\alpha)]$

---

[45] In its usual implementation, this semantics is sometimes called Boolean-valued semantics. Note in this connection that, if there are just two propositions (the true, and the false), then Boolean-valued semantics reduces to two-valued semantics.
[46] For the sake of simplifying our general presentation, we wildly pretend that every singular term is referentially proper. We postpone the technical treatment of improper terms.
[47] Recall that we do not have zero-place function signs.
[48] Recall that we do have zero-place predicates (sentential constants, or subject-less sentences like 'it is raining'), which accordingly denote zero-place propositional functions. A zero-place function from $\mathbb{U}$ into $\mathbb{P}$ is, by definition, simply an element of $\mathbb{P}$.
[49] Direct-quotation gets special treatment, because it is categorially opaque, so the compositional principle cannot apply to it.
[50] For the sake of enormously simplifying the semantic rules, quantifiers are treated *non*-compositionally. If we insist on complete compositionality, then the denotations of expressions must be rendered in a *much* more complicated manner. We opt for simplicity while maintaining compositionality for every other categorial functor.

9.2.	$\upsilon(\alpha\&\beta)$	=	$\upsilon(\&)[\upsilon(\alpha), \upsilon(\beta)]$
9.3.	$\upsilon(\text{that}(\phi))$	=	$\upsilon(\text{that})[\upsilon(\phi)]$
9.4.	$\upsilon(\varepsilon \text{ is true})$	=	$\upsilon(\text{is true})[\upsilon(\varepsilon)]$
9.5.	$\upsilon(\varepsilon_1 \precsim \varepsilon_2)$	=	$\upsilon(\precsim)[\upsilon(\varepsilon_1), \upsilon(\varepsilon_2)]$

36. Evaluating the Logical Expressions

The general categorial scheme only works after we have identified the semantic values of all the primitive logical symbols, both categorial and non-categorial.

36.1. The Non-Categorial Expressions

We begin with the two non-categorial expressions, direct-quotation and quantification.

(r1) $\upsilon(\ulcorner\phi\urcorner) \;=\; \phi$

This merely reflects the idea that the denotation of a direct-quote expression is the literal material within the quotes.

(r2) $\upsilon(\forall\nu\phi) \;=\; glb\{\upsilon'(\phi) : \upsilon' \approx_\nu \upsilon\}$

Here, $\upsilon' \approx_\nu \upsilon$ means that υ' and υ agree on all *simple* expressions except *perhaps* for ν.

36.2. The Categorial Expressions

(r3)	$\upsilon(\ldots\&\ldots)$	=	$\wedge$	[connective]
(r4)	$\upsilon(\sim\ldots)$	=	$\neg$	[connective]
(r5)	$\upsilon(\ldots=\ldots)$	=	$\boldsymbol{I}$	
(r6)	$\upsilon(\ldots\wedge\ldots)$	=	*conj*	[function sign]
(r7)	$\upsilon(\neg\ldots)$	=	*neg*	[function sign]
(r8)	$\upsilon(\text{that}\ldots)$	=	$\text{Id}/\mathbb{P}$	
(r9)	$\upsilon(\ldots\text{is a sentence})$	=	$\mathbb{S}^*$	
(r10)	$\upsilon(\ldots\text{is a proposition})$	=	$\mathbb{P}^*$	
(r11)	$\upsilon(\ldots\text{is true})$	=	$\mathbb{T}$	
(r12)	$\upsilon(\ldots\text{is normal})$	=	$\mathbb{N}$	
(r13)	$\upsilon(\ldots\text{implies}\ldots)$	=	Π	

36.3. The Functions Involved in 36.2

The above restrictions refer to a number of special functions, which we now describe.

36.3.1. $\wedge$ [interprets the conjunction connective]

This is simply the Boolean-meet operation on the algebra of propositions.

36.3.2. $\neg$ [interprets the negation connective]

This is simply the Boolean-complement operation on the algebra of propositions.

36.3.3. *I* [interprets =]

Here, ***I*** is the characteristic function associated with the identity relation on $\mathbb{U}$. In particular,

$\boldsymbol{I}\langle x,y\rangle$	=	1	if	$x\in\mathbb{U}$ & $y\in\mathbb{U}$ & $x=y$
$\boldsymbol{I}\langle x,y\rangle$	=	0		otherwise

In *this* context, please note that 1 is the unit-element of $\mathbb{P}$, and 0 is the zero-element of $\mathbb{P}$.

36.3.4. *conj* [interprets the function sign $\wedge$]

This function maps every ordered pair of formulas to their syntactic conjunction (in order), and maps every ordered pair of propositions to their propositional conjunction (i.e., Boolean meet).

(n1) *conj*: $\mathbb{S}^2\cup\mathbb{P}^2 \rightarrow \mathbb{S}\cup\mathbb{P}$
(n2) $\forall\alpha\beta\in\mathbb{S}$: *conj*$(\alpha,\beta) = (\alpha\&\beta)$
(n3) $\forall pq\in\mathbb{P}$: *conj*$(p,q) = p\wedge q$

36.3.5. *neg*[interprets the function sign $\neg$]

This function maps every formula to its syntactic negation, and maps every proposition to its propositional negation (i.e., Boolean complement).

(n1) *neg*: $\mathbb{S}\cup\mathbb{P} \rightarrow \mathbb{P}$
(n2) $\forall\alpha\in\mathbb{S}$: *neg*$(\alpha) = {\sim}\alpha$
(n3) $\forall p\in\mathbb{P}$: *neg*$(p) = \neg p$

36.3.6. Id/$\mathbb{P}$ [interprets 'that…']

This is the identity function on $\mathbb{P}$:

$$\forall x\in\mathbb{P}: \text{Id}/\mathbb{P}(x) = x.$$

36.3.7. $\mathbb{S}$* [interprets '…is a sentence']

$\mathbb{S}$* is the characteristic function associated with set $\mathbb{S}$. In particular,

$\mathbb{S}*\langle a\rangle = 1$ if $a \in \mathbb{S}$
$\mathbb{S}*\langle a\rangle = 0$ if $a \notin \mathbb{S}$

36.3.8. $\mathbb{P}*$ [interprets '…is a proposition']

$\mathbb{P}*$ is the characteristic function associated with set $\mathbb{P}$. In particular,

$\mathbb{P}*\langle a\rangle = 1$ if $a \in \mathbb{P}$
$\mathbb{P}*\langle a\rangle = 0$ if $a \notin \mathbb{P}$

36.3.9. $\mathbb{T}$ [interprets '…is true']

The function $\mathbb{T}$ divides into two components, according to whether its argument is a sentence or a proposition. In particular:

($\mathbb{T}$1) $\mathbb{T}: \mathbb{S}\cup\mathbb{P} \rightarrow \mathbb{P}$

($\mathbb{T}$2) $\forall\alpha\in\mathbb{S}: \mathbb{T}(\alpha) = \tau(\upsilon(\alpha))$

($\mathbb{T}$3) $\forall p\in\mathbb{P}: \mathbb{T}(p) = p$

The tau-function τ is a primitive function defined on the π-algebra of propositions. (See Part 3).

36.3.10. Π [interprets $\precsim$]

The Π-function is defined as follows.

$\Pi(x,y) = 1$ if $x,y \in \mathbb{S}\cup\mathbb{P}$ & $\mathbb{T}(x) \precsim \mathbb{T}(y)$ [$\precsim$ as defined on π-algebra]
$\Pi(x,y) = 0$ otherwise

37. Examples of Semantic Evaluations

37.1. Example involving a sentence-denoting term

Consider semantically evaluating the formula

$\ulcorner\phi\urcorner$ is a sentence

which is an axiom schema of $\mathcal{T}$. First, we rewrite it as follows.

$\mathbb{S}[\ulcorner\phi\urcorner]$

Next, by the composition principle,

$\upsilon(\mathbb{S}[\ulcorner\phi\urcorner]) =$
$\upsilon(\mathbb{S})\langle\upsilon(\ulcorner\phi\urcorner)\rangle$

$\upsilon(\mathbb{S}) = \mathbb{S}*$
$\upsilon(\ulcorner\phi\urcorner) = \phi$

So:

$$\upsilon(\mathbb{S}[\ulcorner\phi\urcorner]) =$$
$$\upsilon(\mathbb{S})\langle\upsilon(\ulcorner\phi\urcorner)\rangle =$$
$$\mathbb{S}^*\langle\phi\rangle$$

But

$$\mathbb{S}^*\langle\phi\rangle = 1 \quad \text{if} \quad \phi \in \mathbb{S}$$

and, by stipulation, $\phi\in\mathbb{S}$. So $\mathbb{S}^*\langle\phi\rangle = 1$. But $1\in T$, so we conclude that

$$\upsilon \vDash \ulcorner\phi\urcorner \text{ is a sentence}$$

Thus, the schema

$$\ulcorner\phi\urcorner \text{ is a sentence}$$

is, not surprisingly, valid in the proposed semantics.

37.2. Example involving a proposition-denoting term

Consider semantically evaluating the formula

$$\text{that}(\phi) \text{ is true} \leftrightarrow \phi$$

$$\upsilon(\text{that}(\phi) \text{ is true} \leftrightarrow \phi) =$$
$$\upsilon(\text{that}(\phi) \text{ is true}) \leftrightarrow \upsilon(\phi)$$

Now,

$$\upsilon(\text{that}(\phi) \text{ is true}) =$$
$$\upsilon(\text{is true})\langle\upsilon(\text{that}(\phi))\rangle =$$
$$\tau\langle\upsilon(\text{that}(\phi))\rangle$$

And,

$$\upsilon(\text{that}(\phi) =$$
$$\upsilon(\text{that})\langle\upsilon(\phi)\rangle =$$
$$[\text{Id}/\mathbb{P}]\langle\upsilon(\phi)\rangle =$$
$$\upsilon(\phi)$$

so:

$$\upsilon(\text{that}(\phi) \text{ is true} \leftrightarrow \phi) =$$
$$\mathbb{T}\langle\upsilon(\phi)\rangle \leftrightarrow \upsilon(\phi)$$

But $\tau(x)=x$ if $x\in\mathbb{P}$, and by semantic stipulation $\upsilon(\phi)\in\mathbb{P}$, so $\mathbb{T}\langle\upsilon(\phi)\rangle = \upsilon(\phi)$. So,

$$\begin{aligned}
&\upsilon(\text{that}(\phi) \text{ is true } \leftrightarrow \phi) &= \\
&\mathbb{T}\langle \upsilon(\phi)\rangle \leftrightarrow \upsilon(\phi) &= \\
&\upsilon(\phi) \leftrightarrow \upsilon(\phi) &= \\
&1
\end{aligned}$$

Since $1 \in \mathcal{T}$, $\upsilon \vDash [\text{that}(\phi) \text{ is true } \leftrightarrow \phi]$

37.3. Example involving both a sentence-denoting term and a proposition-denoting term

Consider semantically evaluating the formula

$$\ulcorner\phi\urcorner \text{ implies that}(\phi)$$

which is a theorem schema of $\mathcal{T}$. First, we rewrite it as follows.

$$\text{implies}[\ulcorner\phi\urcorner, \text{that}(\phi)]$$

Next, by the composition principle,

$$\begin{aligned}
&\upsilon(\text{implies}[\ulcorner\phi\urcorner, \text{that}(\phi)]) &= \\
&\upsilon(\text{implies})\langle \upsilon(\ulcorner\phi\urcorner), \upsilon(\text{that}(\phi))\rangle
\end{aligned}$$

But

$$\begin{aligned}
\upsilon(\text{implies}) &= \Pi \\
\upsilon(\ulcorner\phi\urcorner) &= \phi \\
\upsilon(\text{that}(\phi)) &= \upsilon(\phi)
\end{aligned}$$

So

$$\begin{aligned}
&\upsilon(\text{implies}[\ulcorner\phi\urcorner, \text{that}(\phi)]) &= \\
&\upsilon(\text{implies})\langle \upsilon(\ulcorner\phi\urcorner), \upsilon(\text{that}(\phi))\rangle &= \\
&\Pi\langle \phi, \upsilon(\phi)\rangle
\end{aligned}$$

Now, Π is defined so that $\Pi(x,y) = 1$ if $x,y \in \mathbb{S}\cup\mathbb{P}$ & $\mathbb{T}(x) \precsim \mathbb{T}(y)$, so

$$\begin{aligned}
&\Pi\langle \phi, \upsilon(\phi)\rangle = 1 \quad \text{if} \\
&\phi \in \mathbb{S}\cup\mathbb{P} \;\&\; \upsilon(\phi) \in \mathbb{S}\cup\mathbb{P} \;\&\; \mathbb{T}(\phi) \precsim \mathbb{T}(\upsilon(\phi))
\end{aligned}$$

Clearly, the first two conjuncts are true. That leaves the third conjunct.

(?) $\quad \mathbb{T}(\phi) \precsim \mathbb{T}(\upsilon(\phi))$

Now, by definition,

($\mathbb{T}$2) $\quad \forall\alpha\in\mathbb{S}: \mathbb{T}(\alpha) = \boldsymbol{\tau}(\upsilon(\alpha))$

($\mathbb{T}$3) $\quad \forall p\in\mathbb{P}: \mathbb{T}(p) = p$

So, since $\phi \in \mathbb{S}$,

$$\mathbb{T}(\phi) \quad = \quad \boldsymbol{\tau}(\upsilon(\phi))$$

and since $\upsilon(\phi) \in \mathbb{P}$,

$$\mathbb{T}(\upsilon(\phi)) \quad = \quad \upsilon(\phi)$$

So the question is whether $\mathbb{T}(\phi) \precsim \mathbb{T}(\upsilon(\phi))$ comes down to the question whether

$$\boldsymbol{\tau}(\upsilon(\phi)) \precsim \upsilon(\phi)$$

But, by definition of the tau-function (Part3), $\boldsymbol{\tau}(x) \precsim x$, for all x in $\mathbb{P}$.

PART 3 – PI-ALGEBRAS

38. Introduction

The semantics proposed for System $\mathcal{T}$ is based on the notion of a pi-algebra[51] (π-algebra), which we explain in this part of the paper. A π-algebra is basically a Boolean algebra augmented by

(1) a one-place function τ
(2) a two-place relation $\precsim$, called a π-ordering.

Whereas the π-ordering is employed in the interpretation of the implication-predicate '$\precsim$', the function τ is used to interpret the truth-predicate.

39. Boolean Algebras

39.1. Generic Boolean Algebras; Finite Conjunction and Disjunction

Most logicians are familiar with Boolean algebras. For the purposes of doing formal semantics as we propose, a *Boolean algebra* is an algebra whose elements are "propositions". This algebra is presumed to be closed under the operations of finite "conjunction" (∧), finite "disjunction" (∨) , and "negation" (¬). A Boolean algebra also has a unit-element 1 (the "necessary proposition") and a zero-element 0 (the "impossible proposition"), presumed to be distinct. These items are required to satisfy the following conditions.

(b1)	$x \wedge x = x$	$x \vee x = x$
(b2)	$x \wedge (y \wedge z) = (x \wedge y) \wedge z$	$x \vee (y \vee z) = (x \vee y) \vee z$
(b3)	$x \wedge y = y \wedge x$	$x \vee y = y \vee x$
(b4)	$x \wedge (x \vee y) = x$	$x \vee (x \wedge y) = x$
(b5)	$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$	$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$
(b6)	$x \wedge \neg x = 0$	$x \vee \neg x = 1$
(b7)	$\neg(x \wedge y) = \neg x \vee \neg y$	$\neg(x \vee y) = \neg x \wedge \neg y$
(b8)	$\neg\neg x = x$	
(b9)	$0 \neq 1$	

A Boolean algebra also admits a two-place "entailment" relation ⩽ satisfying the following conditions.

$$x \leqslant y \quad \leftrightarrow \quad x \vee y = y$$
$$x \leqslant y \quad \leftrightarrow \quad x \wedge y = x$$

The relation ⩽ is a partial ordering, which is to say it satisfies the following conditions.

(p1) $x \leqslant x$
(p2) $x \leqslant y \ \& \ y \leqslant z \quad .\!\rightarrow \quad x \leqslant z$

[51] 'PI' is short for 'para-implication', or 'paradoxical-implication', or 'para-normal-implication', or whatever.

(p3) $x \leqslant y \;\&\; y \leqslant x \;\; .\!\rightarrow \; x = y$

With respect to this ordering,

$$x \wedge y \;=\; glb\{x,y\}$$
$$x \vee y \;=\; lub\{x,y\}$$

Here *glb* is *greatest lower bound*, and *lub* is least upper bound. In other words,

$$z \leqslant x \wedge y \;\leftrightarrow.\; z \leqslant x \;\&\; z \leqslant y$$
$$x \vee y \leqslant z \;\leftrightarrow.\; x \leqslant z \;\&\; y \leqslant z$$

39.2. Complete Boolean Algebras; Infinite Conjunction and Disjunction

Whereas an ordinary Boolean algebra is closed under finite conjunction and disjunction, a *complete Boolean algebra* is closed under infinite conjunction and infinite disjunction. In other words:

$$S \subseteq \mathbb{B} \;\rightarrow\; glb(S) \in \mathbb{B}$$
$$S \subseteq \mathbb{B} \;\rightarrow\; lub(S) \in \mathbb{B}$$

Once again, *glb* is *greatest lower bound*, and *lub* is least upper bound. In other words,

$$x \leqslant glb(S) \;\leftrightarrow\; \forall y\{y \in S \;\rightarrow\; x \leqslant y\}$$
$$lub(S) \leqslant x \;\leftrightarrow\; \forall y\{y \in S \;\rightarrow\; y \leqslant x\}$$

39.3. Filters and Truth

A *filter* on $\mathbb{B}$ is a subset T of $\mathbb{B}$ closed under implication and conjunction. In other words:

$$1 \in T$$
$$p \in T \;\&\; p \leqslant q \;.\!\rightarrow\; q \in T;$$
$$p \in T \;\&\; q \in T \;.\!\rightarrow\; p \wedge q \in T.$$

We use the letter 'T' because a filter may be thought of as the propositions that are ("claimed" to be) true[52]. If you read '$\in T$' as "is true", then the conditions are perfectly obvious. A filter then is a "way of making (some of) the propositions true".

The link between filters and truth is even closer, as given in the following theorem about Boolean algebras.

$$p \leqslant q \;\leftrightarrow\; \text{for every filter } T \text{ on } \mathbb{B}\ \{p \in T \;\rightarrow\; q \in T)$$

In other words,

---

[52] In this connection T can be thought of as a "theory", or a "situation", and a maximal filter can be thought of as a "complete theory" or "possible world".

p implies q
iff
every way of making p true is also a way of making q true

Finally, a filter T is *proper* iff $T \neq \mathbb{B}$, and a filter T is *maximal* if and only iff T is proper and no proper filter properly includes T.

39.4. Complete Filters; Principal Filters

Whereas an ordinary filter on $\mathbb{B}$ is closed under finite conjunction, a *complete filter* is closed under *infinite* conjunction. In other words:

$$S \subseteq T \rightarrow glb(S) \in T.$$

Since $p \wedge q = glb\{p,q\}$, every complete filter is, of course, a filter.

Finally, a complete filter T is *proper* iff $T \neq \mathbb{B}$, and T is *maximal* iff T is proper and no proper complete filter properly includes T.

Intimately tied to the notion of complete filter is the notion of *principal filter*. A filter T is said to be a *principal filter* if and only if T has a least element. Specifically

$$T \text{ is principal } =_{df} \exists x \forall y \{y \in T \leftrightarrow x \leqslant y\}$$

It is easy to show that a filter is complete if and only if it is principal.

39.5. Pi-Ordered Boolean Algebras

If we have no interest in dealing with paradoxical sentences, and the propositions they express, then we can stop at this point, and maintain just one entailment relation, given by $\leqslant$. On the other hand, since we *are* interested in liar sentences, and their ilk, we must augment our logico-algebraic framework to admit the existence of paradoxical sentences (e.g., the liar) and the propositions they express.

The augmentation involves two parts. First, we graft onto a Boolean algebra $\mathbb{B}$ an additional ordering relation $\precsim$, called a pi-ordering. Second, we graft onto $\mathbb{B}$ an additional one-place function τ.

39.6. Pi-Orderings

Let $\mathbb{B}$ be a Boolean algebra. Let $\precsim$ be an relation on $\mathbb{B}$. Then $\precsim$ is a *pi-ordering* (π-ordering) on $\mathbb{B}$ if and only if it satisfies a number of conditions to be gradually presented.

The first condition insists that the pi-ordering $\precsim$ is an extension of the conventional partial-ordering $\leqslant$ on $\mathbb{B}$.

(π1) $p \leqslant q \rightarrow p \precsim q$

Since $\leqslant$ is reflexive, $\precsim$ is reflexive:

(t115) $p \precsim p$

The next condition says that $\precsim$ is transitive.

(π2) $p \precsim q \;\&\; q \precsim r \;.\!\rightarrow\; p \precsim r$

Since $\precsim$ is reflexive and transitive, we can define an affiliated relation, $\approx$, as follows.

(d23) $p \approx q \;=_{df}\; p \precsim q \;\&\; q \precsim p$

It is then easy to prove that $\approx$ is an equivalence relation.

(t116) $p \approx p$

(t117) $p \approx q \rightarrow q \approx p$

(t118) $p \approx q \;\&\; q \approx r \;.\!\rightarrow\; p \approx r$

We also have the requirement that, although the impossible proposition implies the necessary proposition, the converse does not hold. In other words,

(π3) not$[0 \approx 1]$

Next, we add two restrictions that ensure that $\approx$ is not just an equivalence relation, but is moreover a *congruence* relation.

(π4) $p \precsim q \rightarrow \neg q \precsim \neg p$

(π5) $p \precsim q \;\&\; p \precsim r \;.\!\rightarrow\; p \precsim q{\wedge}r$

Indeed this is strengthened as follows.

(π6) $\forall x\{x \in S \rightarrow p \precsim x\} \rightarrow p \precsim \mathit{glb}(S)$

With these principles in hand, one can prove the following.

(c1) $p \approx q \rightarrow \neg p \approx \neg q$

(c2) $p \approx p' \;\&\; q \approx q' \;.\!\rightarrow\; p{\wedge}q \approx p'{\wedge}q'$

Added to the earlier result that $\approx$ is an equivalence relation, these two results yield the theorem that $\approx$ is a congruence relation on the Boolean algebra $\mathbb{B}$.[53]

As can be seen so far, a pi-ordering has many properties in common with $\leqslant$. The main difference is that $\precsim$ is *not presumed* to be anti-symmetric, so distinct propositions can imply each other in the sense of the relation $\precsim$. In fact, this is critical if we wish to model paradoxical sentences.

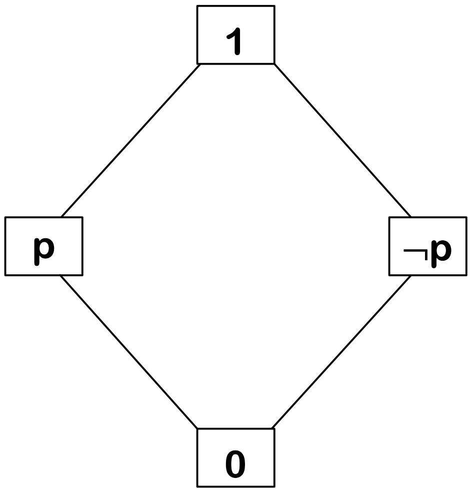
[53] Note carefully that, although $\approx$ is a Boolean congruence, and although $\approx$ is also a congruence with respect to the pi-ordering, $\approx$ will not in general be a congruence with respect to the tau-function; see Section 43. That is why we cannot simply consider the quotient algebra obtained by factoring out $\approx$.

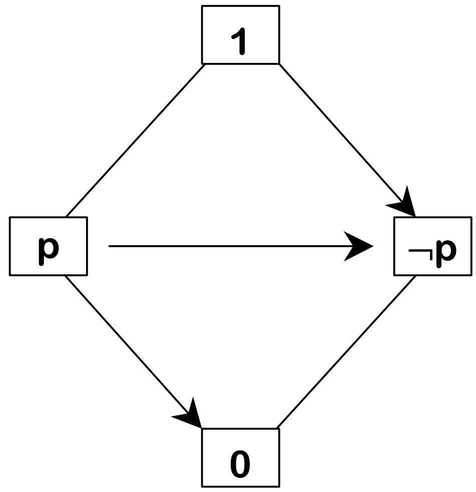
40. Examples of Pi-Ordered Boolean Algebras

Before we continue presenting the restrictions on the pi-ordering, we look at some examples.

40.1. A Simple Example

A simple example of a pi-ordered Boolean algebra is given by taking the four-element Boolean algebra $\mathbb{B}_4$ (left) and superimposing the pi-ordering (right).



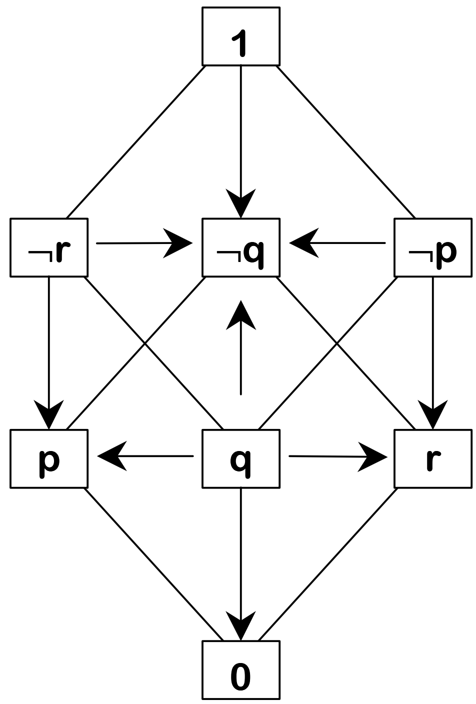


The left diagram, which is a Hasse diagram, explicitly depicts the "covering" relation $\prec$ underlying the partial ordering $\leqslant$. In this example $0\prec p\prec 1$, $0\prec \neg p\prec 1$. The ordering $\leqslant$ is then the smallest partial ordering containing $\prec$. In the right diagram, the non-Boolean implications are represented by the arrows. In this example, there is one *lateral* implication[54] – from p to its negation ¬p. This implication produces two other non-Boolean implications – from p to 0, and from 1 to ¬p. Notice also that $0\leqslant p$, so $0\precsim p$, so $p\approx 0$, and $\neg p\leqslant 1$, so $\neg p\precsim 1$, so $\neg p\approx 1$.

[54] A *lateral* implication, is by definition, any pair (p,q) such that $p\precsim q$ but neither $p\leqslant q$ nor $q\leqslant p$.

40.2. Another Simple Example

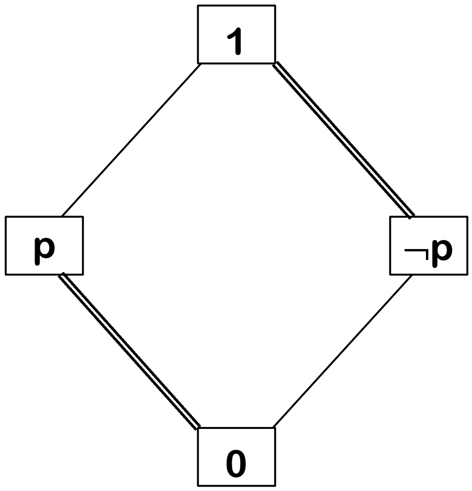
The following is another simple example of a pi-ordered Boolean algebra.

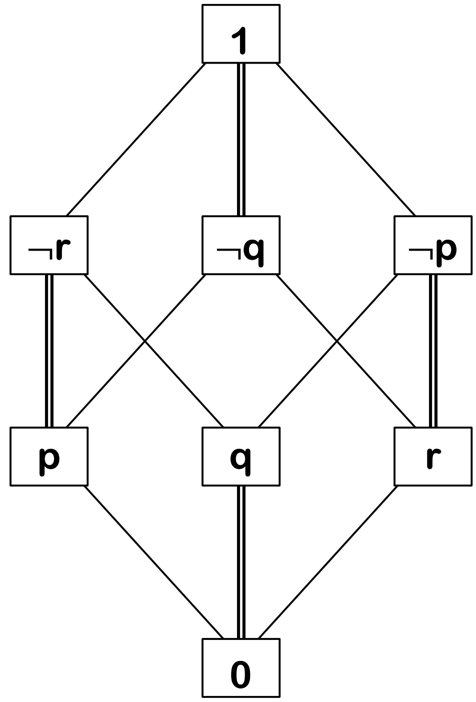


In this example, there are five lateral implications, which in turn produce four other non-Boolean implications.

40.3. An Alternative Diagramming Scheme

Recall that the relation ≈ associated with the pi-ordering ≾ is a congruence relation. This suggests an alternative diagramming scheme that explicitly depicts the relation ≈, and only implicitly depicts the pi-ordering. In the following diagrams, the ≈ relation is depicted along the double lines. From these relations, one can deduce the corresponding ≾-pairs.





41. Pi-Filters

Let ⟨𝔹,≾⟩ be a [complete] pi-ordered Boolean algebra. Then, a [*complete*] *pi-filter* on ⟨𝔹,≾⟩ is, by definition, any subset *T* of 𝔹 satisfying the following conditions.

$1 \in T$

$p \in T \;\&\; p \precsim q \;.\!\rightarrow\; q \in T$

$p \in T \;\&\; q \in T \;.\!\rightarrow\; p \wedge q \in T \;\; [S \subseteq T \rightarrow glb(S) \in T]$

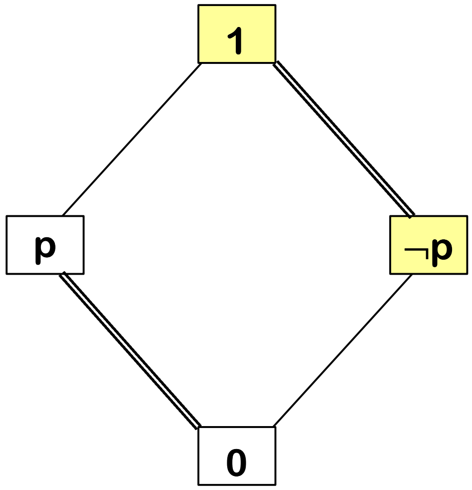
As expected, a [complete] pi-filter *T* is *proper* iff $T \neq 𝔹$, and *T* is *maximal* iff *T* is proper and no proper pi-filter properly includes *T*. Also, a complete pi-filter is a principal filter. In other words,

T is complete $\leftrightarrow \exists p \forall q\{q \in T \leftrightarrow p \leqslant q\}$

41.1. Examples of Maximal Pi-Filters

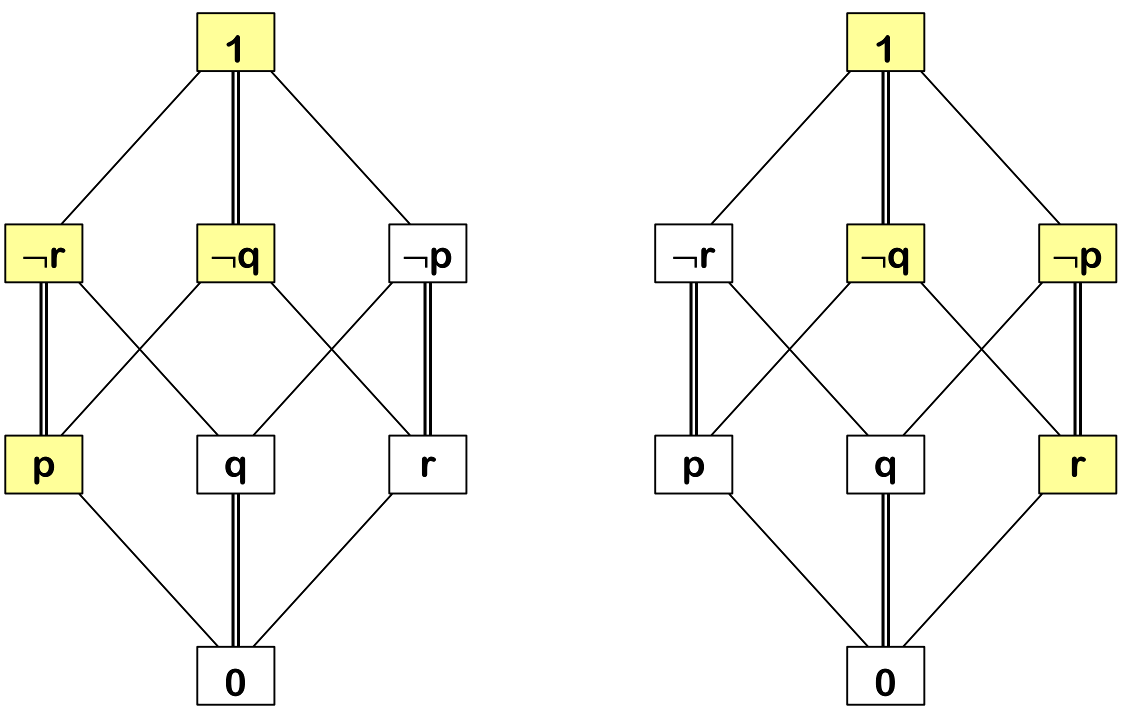
Let us examine our earlier examples of pi-ordered Boolean algebras. The first example gives rise to just one maximal pi-filter, which is depicted in the following by the shaded elements.

Example 1:



The second pi-ordered Boolean algebra gives rise to two maximal pi-filters, depicted as follows. Once again, in each example, the π-filter is the set of shaded elements.

Example 2:



42. Further Restrictions on Pi-Orderings

The remaining restrictions are more complicated, but are nevertheless important.

(π7) There is at least one maximal π-filter on $\langle \mathbb{B}, \precsim \rangle$.

In the semantics for System $\mathcal{T}$, the true propositions are given by a maximal π-filter. This restriction ensures that there is at least one such set of propositions.

(π8) Every maximal π-filter on $\langle \mathbb{B}, \precsim \rangle$ is *prime*.

Here, a prime filter T on a Boolean algebra $\mathbb{B}$ may be defined to be a filter satisfying the following further condition.

(p) $\forall x\{x \in \mathbb{B} \rightarrow. \; x \in T \text{ or } \neg x \in T\}$

This condition ensures that every proposition or its negation is true in the semantics for System $\mathcal{T}$.

The next condition ensures that the implication-relation "meshes" properly with the associated class of maximal π-filters.

(π9) $\forall T\{p \in T \rightarrow q \in T\} \rightarrow p \precsim q$

Here 'T' ranges over *maximal* pi-filters on $\langle \mathbb{B}, \precsim \rangle$.

Notice that our earlier examples satisfy these added restrictions.

43. Tau-Functions

In addition to the pi-ordering, a π-algebra also comes equipped with a one-place operator τ. A tau-function is postulated to satisfy the following conditions.

(τ1) $\tau(1) = 1$

(τ2) $\tau(0) = 0$

(τ3) $\tau(\tau(p)) = \tau(p)$

(τ4) $\tau(p) \leqslant p \rightarrow \tau(p) = p$

(τ5) $p \leqslant \tau(p) \rightarrow \tau(p) = p$

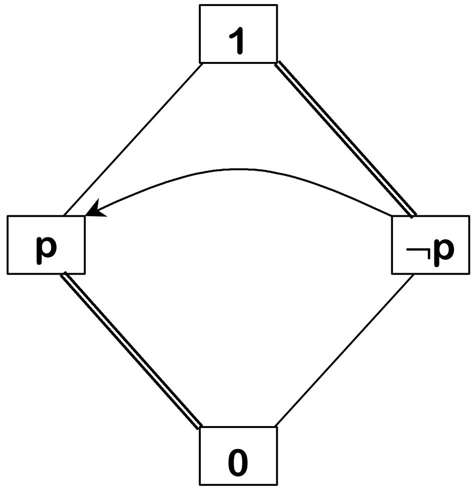
(τ6) $p \leqslant q \rightarrow \tau(p) \leqslant \tau(q)$

(τ7) $\tau(\neg p) = \tau(\neg\tau(p))$

(τ8) $\tau(p) \precsim p$

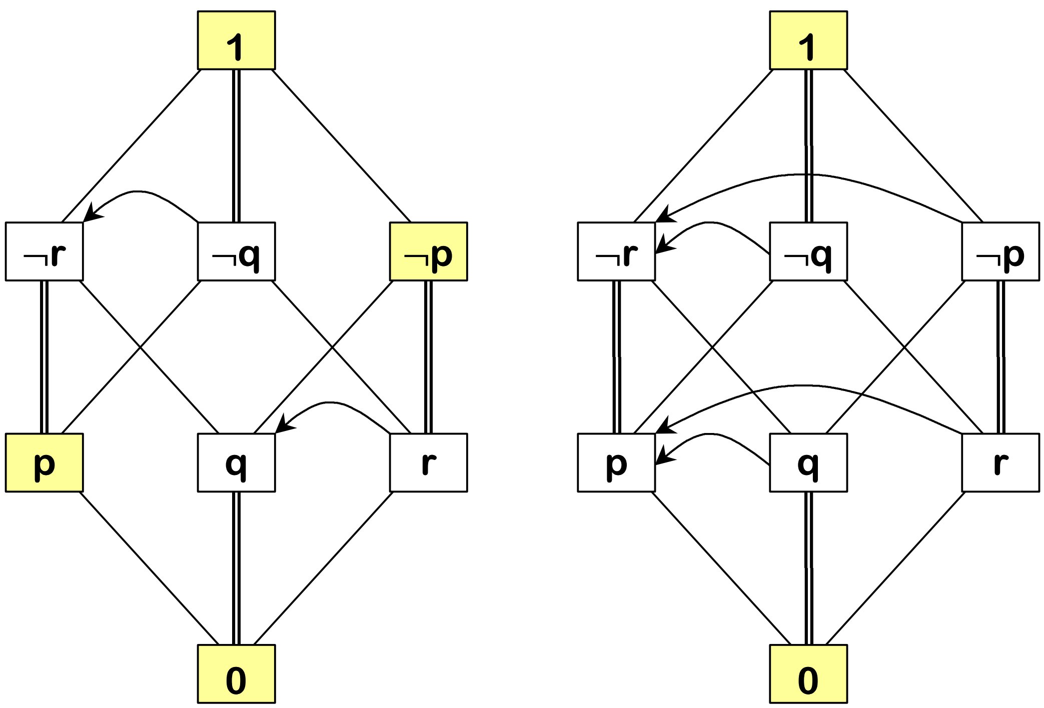
44. Examples of Tau-Functions

Example 1:



In this example, the curved line indicates the *non-trivial* part of the τ-mapping; thus, τ(¬p) = p. The remaining elements are all mapped onto themselves: τ(1) = 1; τ(0) = 0; τ(p) = p.

Examples 2 and 3:



Once again, the curved lines indicate the *non-trivial* part of the τ-mapping; if a mapping is not explicitly indicated, then the element is mapped to itself. [For an explanation of the shaded items, see Section 46.]

45. A Further Restriction on Π-Algebras

There are two more fairly complicated restrictions on pi-algebras. The first one is based on the idea that every paradoxical implication results from a tau-mapping. As a special case, if the tau-mapping is trivial (i.e., for all x, $\tau(x)=x$), then the ordering relations $\leqslant$ and $\precsim$ are identical. This is formally characterized as follows.

(π10) Let $\mathbb{P}$ be a pi-algebra, and let $\approx$ be the associated congruence relation. Define $\cong$ on $\mathbb{P}$ to be the smallest equivalence relation such that: $\tau(x) \cong \tau(x)\wedge x$. Then $\cong$ and $\approx$ are identical.

46. The Normal Core of a Π-Algebra

The π-algebras we are interested in for semantic purposes are further characterized as having a complete “normal core”. This is described as follows.

First, define ‘normal element’ of $\mathbb{P}$ as follows.

(d24) $N[a] \;=_{df}\; \tau(a) = a \;\&\; \tau(\neg a) = \neg a$

(d25) $\mathcal{N} =_{df} \{N[x]: x\in \mathbb{P}\}$

With this in hand, we now state our last requirement on π-algebras.

(π11) $\mathcal{N}$ forms a complete Boolean subalgebra of $\mathbb{P}$.

This means in particular that the following are satisfied.

(1) $S\subseteq\mathcal{N} \;\rightarrow\; glb(S) \in \mathcal{N}$
$a\in\mathcal{N} \;\&\; b\in\mathcal{N} \;.\!\rightarrow\; a\wedge b \in \mathcal{N}$

(2) $S\subseteq\mathcal{N} \;\rightarrow\; lub(S) \in \mathcal{N}$
$a\in\mathcal{N} \;\&\; b\in\mathcal{N} \;.\!\rightarrow\; a\vee b \in \mathcal{N}$

Examples:

In each of the previous two examples, the normal core is depicted by the shaded elements.