THE REVIEW OF SYMBOLIC LOGIC
Volume 12, Number 2, June 2019

A UNIFIED THEORY OF TRUTH AND PARADOX

LORENZO ROSSI
Department of Philosophy (KGW), University of Salzburg

Abstract. The sentences employed in semantic paradoxes display a wide range of semantic behaviours. However, the main theories of truth currently available either fail to provide a theory of paradox altogether, or can only account for some paradoxical phenomena by resorting to multiple interpretations of the language, as in (Kripke, 1975). In this article, I explore the wide range of semantic behaviours displayed by paradoxical sentences, and I develop a unified theory of *truth and paradox*, that is a theory of truth that also provides a unified account of paradoxical sentences. The theory I propose here yields a threefold classification of paradoxical sentences—liar-like sentences, truth-teller–like sentences, and revenge sentences. Unlike existing treatments of semantic paradox, the theory put forward in this article yields a way of interpreting all three kinds of paradoxical sentences, as well as unparadoxical sentences, within a single model.

§1. Introduction. Semantic predicates such as truth, satisfaction, and denotation play a crucial role in several contemporary theories of meaning.[1] However, semantic predicates are famously problematic: simple and intuitive assumptions about the principles governing them, together with a *modicum* of logic and syntax theory, yield well-known paradoxes. Consider the self-applicable predicate '... is true'. An apparently compelling intuition suggests that it should obey the following informal principle:

(NAÏVETÉ) For every sentence φ, φ and "φ' is true' are equivalent.

But now consider a liar sentence λ equivalent to "λ' is not true'. If the truth predicate obeys NAÏVETÉ, then λ is true if and only if λ is not true, a contradiction. But (in classical logic, and several other logics) everything follows from a contradiction, whereby every sentence is true.

The Liar Paradox is not an isolated phenomenon. Semantic notions can be used to form several kinds of *sensu lato* paradoxical sentences, which display a wide range of semantic behaviours. For instance, a *truth-teller* sentence τ equivalent to "τ' is true' can be consistently validated, falsified, or assigned any other semantic value by any semantics compatible with NAÏVETÉ. *Revenge paradoxes* show that certain semantic notions, related to naïve truth, are inexpressible in a target theory. Analogous paradoxes arise for satisfaction, denotation, and other semantic notions. For simplicity, I focus on the truth predicate

Received: December 15, 2014.
2010 *Mathematics Subject Classification*: 03B50.
Key words and phrases: semantic paradoxes, truth predicate, naïve truth, non-classical logics, revenge paradoxes.

[1] Consider e.g., Davidsonian and Montagovian semantics (see e.g., Davidson (1967) and Montague (1974), Chierchia & McCconnell-Ginet (2000), respectively), or truthmaker semantics Fine (2017).

doi:10.1017/S1755020319000078

and on relatively simple languages, which are however expressive enough to encode some basic syntactic mechanisms.[2]

The aim of this article is to investigate the semantics of sentences involving the truth predicate, including liar sentences, truth-teller sentences, revenge sentences, and the like. More specifically, the article provides a theory of truth that also accounts for and classifies all the paradoxical sentences involving truth. The motivation behind the theory offered here is that, if a semantics for a natural language employs a self-applicable truth predicate, then that semantics is going to have to interpret all kinds of sentences involving the truth predicate, whether they are 'unproblematic', or in some sense 'paradoxical'. To my knowledge, the modern systematic analysis of paradoxes was initiated by *fixed-point* theories of truth (Kripke, 1975) and *revision* theories of truth (Gupta, 1982; Herzberger, 1982a, 1982b).[3] Recent years have seen a growth of *graph-theoretical* approaches, which are very successful at identifying structural features of paradoxical sentences.[4] Nevertheless, the theories of truth and paradox currently available do not seem to provide a unified semantics for paradoxical and nonparadoxical sentences. For one thing, existing approaches resort to several models to account for the semantics of certain paradoxical sentences—this includes Kripke's approach and the Hezberger–Gupta approach. More precisely, existing theories cannot provide an interpretation of unproblematic sentences, such as 'snow is white' or $t = t$, and of all the various kinds of paradoxical sentences within one single model.[5] For another, existing semantic theories typically do not treat revenge-paradoxical sentences, since adding revenge-breeding notions to their target language would make them trivial. Revenge-breeding notions are therefore also relegated to the meta-theory.[6]

In this article, I propose a unified theory of truth and paradox, i.e., a single model that interprets both nonparadoxical and paradoxical sentences. The interpretation of paradoxical sentences consists in assigning them special semantic values, encoding (as much as possible) their semantic behaviour. The theory I develop differentiates between three main kinds of paradoxical cases: liar-like sentences, truth-teller–like sentences, and revenge sentences. As I will argue, these cases exhaust all the main paradoxes of truth. More precisely, the theory I present here can accommodate any compositional interpretation of the logical vocabulary, without affecting the resulting classification of paradoxes. The proposed classification of paradoxes is therefore robust, and should be shared by any compositional approach to naïve truth.[7]

---

[2] The theory I am going to propose can be easily extended to other semantic predicates and richer languages. See McGee (1991, pp. 31–37), for details on the relations between truth, satisfaction, and denotation.

[3] See also Burgess (1986, 1988), Gupta & Belnap (1993), Kremer (2009).

[4] See e.g., Davis (1979), Hazen (1981), Barwise & Etchemendy (1987), Gaifman (1988), Gaifman (1992), Yi (1999), Gaifman (2000), Cook (2004), Maudlin (2004), Cook (2006), Schlenker (2007), Walicki (2009), Rabern, Rabern, & Macauley (2013), Cook (2014), Dyrkolbotn & Walicki (2014), Hansen (2015), Walicki (2017), Beringer & Schindler (2017).

[5] Consider the treatment of liar and truth-teller sentences in fixed-point and revision theories: in both approaches, several models have to be considered to differentiate liar sentences, truth-teller sentences, and truths or falsities of the base language.

[6] For discussion, see Field (2007), Leitgeb (2007), Rossi (2019).

[7] The interest of the study of paradoxes goes well beyond the goal of interpreting the sentences of a language featuring a self-applicable truth predicate. The study of paradoxes has led to discover new limitative results and to determine which logical principles and evaluation schemes are compatible with NAÏVETÉ (see e.g., Kripke (1975), Friedman & Sheard (1987), McGee (1985), Restall (1992), Hájek, Paris, & Shepherdson (2000), Field (2002, 2003), Halbach,

The plan of the article is as follows. In §2, I present some representative semantic paradoxes, and I explore the challenges their account poses to a semantic theory of truth. I argue that the kinds of paradoxes presented in §2 yield an exhaustive taxonomy of semantic antinomies. In §3, I provide some heuristics. In §4 I develop the proposed theory of truth and paradox. Technically, this theory employs a combination of graph-theoretic tools, fixed-point constructions, and revision sequences. I argue that the proposed theory satisfactorily accounts for the semantics of the paradoxical sentences classified in §2, and I sketch some prospects for further developments. §5 concludes. The main proofs are given in the Appendix.

§2. Naïveté and paradoxes. Naïveté about truth—the idea that φ and "'φ' is true' are equivalent—can be made precise in a number of ways. For present purposes, I characterise it in semantic terms. An *evaluation* is a function from the sentences of the target language to a nonempty value space V, including some designated values. Let $\mathsf{Tr}(\ulcorner \varphi \urcorner)$ abbreviate "'φ' is true', where $\ulcorner \varphi \urcorner$ is a name of φ. At first approximation, an evaluation e is said to be *naïve* if one of the following requirements is satisfied:

– (INTERSUBSTITUTIVITY) $\qquad e(\varphi) = e(\varphi^{\mathsf{Tr}})$
  where φ^{Tr} is the result of substituting (possibly nonuniformly) a subformula ψ of φ with $\mathsf{Tr}(\ulcorner \psi \urcorner)$ or *vice versa*.

– (T-SCHEMA) $\qquad e(\mathsf{Tr}(\ulcorner \varphi \urcorner) \leftrightarrow \varphi) = \mathbf{d}$
  where $\mathbf{d}$ is a designated value of V.

I generically speak of NAÏVETÉ when it makes no difference whether an evaluation function satisfies INTERSUBSTITUTIVITY or the T-SCHEMA. I now briefly present some important paradoxical sentences, outlining the challenges that capturing their semantics poses.

2.1. Liar-like sentences. The Liar Paradox features a sentence that, roughly speaking, says that that very sentence is not true. For instance, consider the following sentence:

(λ) The sentence labelled with '(λ)' is not true.

Liar sentences can be used to show that no *classical* evaluation satisfies NAÏVETÉ. For suppose that a classical evaluation e satisfies NAÏVETÉ, let λ be the sentence $\neg\mathsf{Tr}(\ulcorner \lambda \urcorner)$,

Leitgeb, & Welch (2003), Halbach & Horsten (2006), Priest (2006), Cieśliński (2007), Field (2008), Beall (2009), Horsten (2009), Zardini (2011), Cobreros, Égré, Ripley, & van Rooij (2013), Field (2013), Nicolai & Rossi (2018), Murzi & Rossi (2019)). Moreover, the analysis of paradoxes has been instrumental to determine the expressive power of theories of truth (see e.g., Ketland (2003), Beall (2006, 2007a, 2007b), Cook (2007), Field (2007), Leitgeb (2007), Maudlin (2007), Priest (2007), Restall (2007), Scharp (2007), Simmons (2007), Shapiro (2011), Scharp (2013), Rossi (2019)). Finally, the investigation of semantic paradoxes has revealed connections between theories of truth and questions concerning coding, circularity, self-reference, and nonwell-foundedness (see e.g., Yablo (1985), Gaifman (1988), McCarthy (1988), Visser (1989), Gaifman (1992), Yablo (1993), Priest (1997), Yi (1999), Gaifman (2000), Beall (2001), Leitgeb (2002), Bueno & Colyvan (2003), Ketland (2004), Cook (2004, 2006), Yablo (2006), Schlenker (2007), Cook (2014), Halbach & Visser (2014a, 2014b), Beringer & Schindler (2017)).

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

and consider the value of $e(\neg\mathsf{Tr}(\ulcorner\lambda\urcorner))$. Since e is classical, V consists of two semantic values, $\mathbf{1}$ and $\mathbf{0}$. And since e is classical, either $e(\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)) = \mathbf{1}$ or $e(\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)) = \mathbf{0}$. However, if $e(\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)) = \mathbf{1}$ then, by the classical semantics for negation, NAÏVETÉ, and the definition of λ, $e(\mathsf{Tr}(\ulcorner\lambda\urcorner)) = \mathbf{0} = e(\neg\mathsf{Tr}(\ulcorner\lambda\urcorner))$, which is impossible. Similarly, if $e(\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)) = \mathbf{0}$, then $e(\mathsf{Tr}(\ulcorner\lambda\urcorner)) = \mathbf{1} = e(\neg\mathsf{Tr}(\ulcorner\lambda\urcorner))$, which is equally impossible.

The same conclusion is reached with other paradoxical sentences: Curry's Paradox employs a sentence κ identical to the sentence 'if 'κ' is true, then $\perp$' (where $\perp$ is some conventional falsity); McGee (1985)'s Paradox employs a sentence μ identical to 'not every iteration of Tr in front of 'μ' is true', and both can be used to show that classical evaluations do not satisfy NAÏVETÉ.

The Liar Paradox, Curry's Paradox, McGee's Paradox and many others arguably show that, in order to interpret a language with a naïve truth predicate, a nonclassical semantics is required. In order to accommodate NAÏVETÉ, several nonclassical semantics expand the value space V with intermediate values between $\mathbf{0}$ and $\mathbf{1}$, generalising the evaluation clauses accordingly. In this way, the sentences that receive a classical value ($\mathbf{0}$ or $\mathbf{1}$) obey the principles of classical logic, while the sentences that are assigned a nonclassical value display a different semantic behaviour. I clarify this point with an example (which will be useful later).

A *partial evaluation* is a function that assigns to the sentences of the target language one amongst the values $\mathbf{1}$, $\mathbf{0}$, and $\mathbf{1/2}$, and that satisfies the following criteria:[8]

- The value of $\neg\varphi$ is $\mathbf{1}$ minus the value of φ.
- The value of $\varphi \wedge \psi$ is the minimum of the values of φ and ψ.
- The value of $\forall x\varphi$ is the infimum of the values of its instances $\varphi(t)$.

Several semantics for naïve truth are based on partial evaluations.[9] Liar, Curry, and McGee sentences can be assigned value $\mathbf{1/2}$ by partial evaluations, together with their negation.

2.2. Truth-teller sentences. While the Liar Paradox rules out some evaluations for naïve truth, the Truth-teller Paradox presents quite an opposite scenario. The paradox involves a sentence that, roughly, says that that very sentence is true, e.g.:

(τ) The sentence labelled with '(τ)' is not true.

Let τ be the sentence $\mathsf{Tr}(\ulcorner\tau\urcorner)$. No feature of τ 'forces' one value assignment over another, unlike liar sentences which are forced by NAÏVETÉ to have the same value as their negation.

The fact that truth-teller sentences can be assigned any available value might make them appear to be unproblematic, but this is far from being the case. In most semantic theories of truth, truth-teller sentences are assigned a semantic value—be it $\mathbf{1}$, $\mathbf{0}$, $\mathbf{1/2}$, or another intermediate value—exactly as any other sentence, like $\forall x(x = x)$ or λ. But assigning value $\mathbf{1}$ to $\forall x(x = x)$ seems appropriate, for few will doubt of the truths of the theory of identity. And assigning value $\mathbf{1/2}$ to λ also seems appropriate, for NAÏVETÉ forces λ to have the same value of $\neg\lambda$, showing that no classical value is appropriate for liar sentences.

8 For more on partial evaluations, see Kleene (1952, chap. XII) and Blamey (2002).

9 Examples include strong Kleene semantics (see e.g., Kripke (1975)), the logic of paradox (see e.g., Asenjo (1966), Priest (1979)), strict-tolerant and tolerant-strict semantics (see e.g., Cobreros, Égré, Ripley, & van Rooij (2012), Nicolai & Rossi (2018)).

However, no 'standard' value seems appropriate for $\mathsf{Tr}(\ulcorner\tau\urcorner)$, because no such value seems to be 'the' right one for τ, in that there are no grounds for choosing a value over another. No 'standard' value (such as $\mathbf{1}$, $\mathbf{0}$ or $\mathbf{1/2}$) captures the fact that truth-teller sentences *can* be assigned *any* value, thus suitably representing their semantic behaviour. Several theories of truth can only account for this behaviour resorting to *multiple models*: for instance, both fixed-point and revision theories of truth capture the difference between liar and truth-teller sentences by showing that the latter can be assigned several values or revision sequences, unlike the former.[10]

2.3. Revenge sentences. Even though truth-teller sentences are puzzling, they are relatively inoffensive as they do not yield inexpressibility results.[11] *Revenge sentences* are much less innocuous. Revenge paradoxes are arguments to the effect that certain semantic notions, related to naïve truth, are not expressible in a target theory. Consider again the treatment of liar sentences in semantics based on partial evaluations (and in which $\mathbf{1}$ is the only designated value). In such theories, sentences such as λ are assigned value $\mathbf{1/2}$. Despite the fact that liar sentences do not receive value $\mathbf{1}$, they fail to be declared 'not true', since their negation receives also value $\mathbf{1/2}$, and not $\mathbf{1}$. In order to properly express the fact that liar sentences fail to be true, one could employ a notion of *determinateness* that maps both values $\mathbf{0}$ and $\mathbf{1/2}$ to $\mathbf{1}$. Consider therefore a unary operator D, with the following semantics (for an evaluation function e):

$$e(\mathsf{D}(\varphi)) = \begin{cases} \mathbf{1}, & \text{if } e(\varphi) = \mathbf{1}, \\ \mathbf{0}, & \text{if } e(\varphi) \neq \mathbf{1}. \end{cases}$$

Using D, it should be possible to declare liar sentences 'not determinate', assigning value $\mathbf{1}$ to $\neg\mathsf{D}(\lambda)$. However, such an operator D is inexpressible in the setting we assumed. For suppose otherwise, and consider the sentence λ_{d} identical to $\neg\mathsf{D}(\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner))$. If $e(\lambda_{\mathsf{d}}) = \mathbf{1}$, then $e(\mathsf{D}(\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner))) = \mathbf{1} = e(\neg\mathsf{D}(\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner)))$, which is impossible. Similarly, if $e(\lambda_{\mathsf{d}}) = \mathbf{0}$, then $e(\mathsf{D}(\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner))) = \mathbf{0} = e(\neg\mathsf{D}(\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner)))$, which is impossible as well. We have a revenge paradox: (bivalent) determinateness is inexpressible.

Revenge paradoxes pose a serious threat to theories of truth: if successful, they show that their target theories have severe expressive limitations. Proponents of revenge-prone theories of truth have typically sought to avoid the problem by arguing that the revenge-

10 Albert Visser (1984, secs. 3.4–3.5) argues that some four-valued models can distinguish between liar sentences and truth-teller sentences. As he puts it: '[o]ne attractive feature of four valued logic for the study of the Liar Paradox is the possibility of making certain distinctions *within one single model*. [. . .] I present various models in which the Liar is both true and false and the [truth-teller] neither true nor false. The intuitive idea here is that the Liar *must* be true, *must* be false; the [truth-teller] *need not* be true, *need not* be false.' Visser (1984, pp. 181–182). Nevertheless, it is not obvious that 'neither true nor false' has a better claim to capture the semantic behaviour of truth-teller sentences than the other values in four valued semantics. If 'true' and 'false' (understood as semantic values) are not to be assigned to the truth-teller because it '*need not* be true, *need not* be false', the same could be said of 'neither true nor false' itself, since truth-teller sentences *need not* be neither true nor false either. This reasoning generalises easily to any 'standard' semantic value.

11 Mortensen & Priest (1981) and Billon (2013) argue that the peculiar semantic behaviour of truth-teller sentences can be turned into a proper contradiction (see also Smith (1984)). For more discussion, see Sorensen (2001, chap. 11) and Greenough (2011).

paradoxical notions are not genuine semantic notions.[12] And since revenge-breeding notions are not expressible in the theories they are directed against, existing theories of truth simply do not consider revenge sentences. However, it is very unclear whether there are principled reasons to reject notions such as bivalent determinateness, while keeping naïve truth and other notions that breed 'standard' paradoxes. If at least some revenge-paradoxical notions are genuine semantic notions, a theory of truth and paradox needs to interpret them as well.[13]

2.4. A complete picture. The paradoxes described in §2.1–§2.3 can be be plausibly argued to cover the main kinds of semantic behaviours that are relevant to a theory of truth and paradox—that is, those that require different kinds of semantic value assignments. The idea, roughly, is the following.

Suppose a semantics for the logical vocabulary and a value space have been selected, and consider an arbitrary sentence φ. There are two possibilities: φ is either compatible with exactly one semantic value assignment, or it isn't. In the former case, φ has either a classical value (as in the case of $\forall x(x = x)$), or a nonclassical value (as in the case of λ). In the latter case, φ is either compatible with more than one semantic value (as in the case of τ), or it is not compatible with any semantic value (as in the case of λ_{d}).

More systematically, this suggests the following classification:

- Sentences that are compatible with exactly one semantic value:
  - *Nonparadoxical sentences*: they are assigned a classical semantic value.
  - *Liar-like sentences*: they are assigned a nonclassical semantic value.
- Sentences that are not compatible with exactly one semantic value:
  - *Truth-teller–like sentences*: they are assigned a special semantic value that indicates that they are compatible with more than one (standard) semantic value.
  - *Revenge sentences*: they are assigned a special semantic value that indicates that they are incompatible with every (standard) semantic value.

The above taxonomy arguably covers all the possible outcomes of an evaluation of φ. In the next section, I develop a theory that incorporates and makes explicit all the above cases, thus yielding a theory of naïve truth as well as an account of paradoxical sentences.

§3. Heuristics. In this section, I provide some heuristics for the theory to be developed in §4. I show how certain kinds of graphs, called *semantic graphs*, can be used to decompose sentences and assign them semantic values, exemplifying this process with nonpara-

---

[12] See e.g., Priest (2006), Field (2007), Beall (2009).

[13] For arguments for the legitimacy of revenge-paradoxical notions, see e.g., Cook (2007), Leitgeb (2007), Scharp (2007, 2013), Rossi (2019), Murzi & Rossi (2019). The indefinite extensibility approach developed by Cook (2007, 2009), Cook & Tourville (2016) and Schlenker (2010) does not refrain from interpreting revenge sentences. In a nutshell, in this approach semantic paradoxes are interpreted in an indefinitely extensible succession of evaluations, with an indefinitely extensible collection of semantic values. Nevertheless, due to its use of an indefinitely extensible collection of values, this approach also resorts to infinitely many (actually, nonset-sized many) models in order to characterise certain semantic paradoxes (more specifically, the phenomenon of revenge).

doxical sentences (§3.1), liar-like sentences (§3.2), truth-teller–like sentences (§3.3), and revenge sentences (§3.4).

My starting point is a very basic question: what information is needed to evaluate a sentence φ? The answer depends on the logical form of φ. If φ is an atomic sentence of the base (i.e., truth-predicate-free) language, its value is determined by the selected model of the base language. For instance, if φ is $P(t_0, \ldots, t_n)$, its value is determined by whether the individuals denoted by $t_0, \ldots, t_n$ (in the selected model) are in the extension of the predicate $P(x_0, \ldots, x_n)$ (in the selected model). If φ is a logically complex sentence, compositionality dictates that the value of φ depend on the immediate subsentences of φ. For instance, if φ is $\neg\psi$, the value of φ depends on the value of ψ; if φ is $\psi \wedge \chi$, the value of φ depends on the values of ψ and χ, and so on. NAÏVETÉ suggests how to extend this compositional picture to the truth predicate: the value of $\mathsf{Tr}(\ulcorner \psi \urcorner)$ depends on the value of ψ, even though ψ is not a subsentence of $\mathsf{Tr}(\ulcorner \psi \urcorner)$.[14]

With this informal picture of semantic value assignments in mind, I look at the main kinds of nonparadoxical or paradoxical sentences discussed above.

3.1. Nonparadoxical sentences. Consider the sentence $\mathsf{Tr}(\ulcorner t = t \urcorner)$. In order to assign it a value, one needs the value of $t = t$. I represent this process via a downward arrow, in a suitable labelled graph:

$$\begin{tikzcd}
\mathsf{Tr}(\ulcorner t = t \urcorner) \arrow[d] \\
t = t
\end{tikzcd}$$

Fig. 1. Graph for $\mathsf{Tr}(\ulcorner t = t \urcorner)$.

The semantic value of $t = t$ is unproblematic: being an atomic formula of the base language, its semantic value is determined by the selected model of the base language, and it is clearly $\mathbf{1}$. Once $t = t$ is assigned value $\mathbf{1}$, NAÏVETÉ suggests to assign the same value to $\mathsf{Tr}(\ulcorner t = t \urcorner)$.

This process easily handles more complex sentences. Consider $\mathsf{Tr}(\ulcorner t = t \urcorner) \leftrightarrow t = t$, that is $(\mathsf{Tr}(\ulcorner t = t \urcorner) \rightarrow t = t) \wedge (t = t \rightarrow \mathsf{Tr}(\ulcorner t = t \urcorner))$. Decomposing it iteratively, following the intuition outlined above, the following graph obtains:

$$\begin{tikzcd}
 & & & (\mathsf{Tr}(\ulcorner t = t \urcorner) \rightarrow t = t) \wedge (t = t \rightarrow \mathsf{Tr}(\ulcorner t = t \urcorner)) \arrow[dll] \arrow[drr] & & & \\
 & \mathsf{Tr}(\ulcorner t = t \urcorner) \rightarrow t = t \arrow[dl] \arrow[dr] & & & & t = t \rightarrow \mathsf{Tr}(\ulcorner t = t \urcorner) \arrow[dl] \arrow[dr] & \\
\mathsf{Tr}(\ulcorner t = t \urcorner) \arrow[d] & & t = t & & t = t & & \mathsf{Tr}(\ulcorner t = t \urcorner) \arrow[d] \\
t = t & & & & & & t = t
\end{tikzcd}$$

Fig. 2. Graph for $\mathsf{Tr}(\ulcorner t = t \urcorner) \leftrightarrow t = t$.

[14] Versions of the approach to evaluation just sketched can be found, e.g., in Kripke (1975), Yablo (1982), and Leitgeb (2005) amongst others.

$t = t$ cannot be decomposed any further, so the graph-formation process stops, and values can be assigned. First, $t = t$ has value $\mathbf{1}$. Since $t = t$ has value $\mathbf{1}$, also $\mathsf{Tr}(\ulcorner t = t \urcorner)$ has value $\mathbf{1}$. But then also $(\mathsf{Tr}(\ulcorner t = t \urcorner) \to t = t)$ and $(t = t \to \mathsf{Tr}(\ulcorner t = t \urcorner))$ have value $\mathbf{1}$, because they are conditionals whose antecedents and consequents have value $\mathbf{1}$. Finally, $(\mathsf{Tr}(\ulcorner t = t \urcorner) \to t = t) \wedge (t = t \to \mathsf{Tr}(\ulcorner t = t \urcorner))$, a conjunction whose conjuncts have value $\mathbf{1}$, has value $\mathbf{1}$ as well.

3.2. Liar-like sentences. Consider now a sentence λ identical to $\neg\mathsf{Tr}(\ulcorner \lambda \urcorner)$. It is a negated formula, so in order to evaluate it one should look at $\mathsf{Tr}(\ulcorner \lambda \urcorner)$:

$$\begin{tikzcd} \neg\mathsf{Tr}(\ulcorner \lambda \urcorner) \arrow[d] \\ \mathsf{Tr}(\ulcorner \lambda \urcorner) \end{tikzcd}$$

And in order to evaluate $\mathsf{Tr}(\ulcorner \lambda \urcorner)$, one should evaluate the sentence resulting by disquotationally eliminating Tr from $\mathsf{Tr}(\ulcorner \lambda \urcorner)$. But this sentence is $\neg\mathsf{Tr}(\ulcorner \lambda \urcorner)$ itself. A *loop* results:

$$\begin{tikzcd} \neg\mathsf{Tr}(\ulcorner \lambda \urcorner) \arrow[d] \\ \mathsf{Tr}(\ulcorner \lambda \urcorner) \arrow[u, bend left] \end{tikzcd}$$

Fig. 3. Graph for the liar sentence $\neg\mathsf{Tr}(\ulcorner \lambda \urcorner)$.

Recognisably, the search for the information that is required to evaluate $\neg\mathsf{Tr}(\ulcorner \lambda \urcorner)$ is finished: in order to evaluate $\neg\mathsf{Tr}(\ulcorner \lambda \urcorner)$ one needs the value of $\mathsf{Tr}(\ulcorner \lambda \urcorner)$, and in order to evaluate the latter one needs the value of the former. Can this information be used to assign values?

In fact it can. Even if the above graph does not bottom out in sentences of the base language, it provides information about the relation between $\neg\mathsf{Tr}(\ulcorner \lambda \urcorner)$ and $\mathsf{Tr}(\ulcorner \lambda \urcorner)$ that can lead to a value assignment once a semantics for the logical vocabulary has been selected. Suppose we adopt the following semantics for negation:[15]

$$e(\neg\varphi) = \mathbf{1} - e(\varphi).$$

The above graph, in combination with the above clause for negation, indicates that a constraint should be placed on the possible values of $\neg\mathsf{Tr}(\ulcorner \lambda \urcorner)$, which can be written as an equation:

$$\text{the value of } \neg\mathsf{Tr}(\ulcorner \lambda \urcorner) = \mathbf{1} - \text{ the value of } \mathsf{Tr}(\ulcorner \lambda \urcorner).$$

Moreover, the informal method followed so far employs the following evaluation clause for truth attributions (which derives from NAÏVETÉ):

[15] Where e is an evaluation whose range includes $\mathbf{1}$, and on which subtraction is well-defined.

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

$$e(\mathsf{Tr}(\ulcorner\varphi\urcorner)) = e(\varphi)$$

and this also suggests a constraint on the possible values of $\mathsf{Tr}(\ulcorner\lambda\urcorner)$, that is the following equation:

$$\text{the value of } \mathsf{Tr}(\ulcorner\lambda\urcorner) = \text{the value of } \neg\mathsf{Tr}(\ulcorner\lambda\urcorner).$$

All the relational information provided by the selected semantics and graph 3 has been associated with the sentences appearing in the graph. Such information determines an equation system which expresses simultaneous constraints on the possible values of $\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)$ and $\mathsf{Tr}(\ulcorner\lambda\urcorner)$:

$$x = \mathbf{1} - y$$
$$y = x.$$

It can therefore be checked whether such constraints can be univocally satisfied, i.e., whether the system has a unique solution. In this case, yes: $x = \mathbf{1/2} = y$ (if $\mathbf{1/2}$ is in the chosen value space).[16]

It is easy to see that the analysis of Curry's or McGee's sentences yields similar outcomes. Consider a Curry sentence κ identical to $\mathsf{Tr}(\ulcorner\kappa\urcorner) \to \bot$. Its graph is as follows:

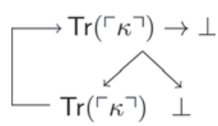

Fig. 4. Graph for the Curry sentence $\mathsf{Tr}(\ulcorner\kappa\urcorner) \to \bot$.

Clearly $\bot$ is assigned value $\mathbf{0}$, while the remaining two sentences in graph 4 yield the following system (interpreting $\to$ as a material conditional):

$$\text{the value of } \mathsf{Tr}(\ulcorner\kappa\urcorner) \to \bot = \max[\mathbf{1} - \text{the value of } \mathsf{Tr}(\ulcorner\kappa\urcorner), \mathbf{0}]$$
$$\text{value of } \mathsf{Tr}(\ulcorner\kappa\urcorner) = \text{the value of } \mathsf{Tr}(\ulcorner\kappa\urcorner) \to \bot.$$

Re-writing it with variables we obtain the following system

$$x = \min[\mathbf{1} - y, \mathbf{0}]$$
$$y = x$$

which has a unique solution, again $x = \mathbf{1/2} = y$.

Finally, consider a McGee sentence μ identical to $\neg\forall n\mathsf{Tr}(\ulcorner \underset{\cdot}{g}(n, \ulcorner\mu\urcorner)\urcorner)$, where $\underset{\cdot}{g}$ is the (object-linguistic term representing the primitive recursive) function g such that:

$$g(n, \varphi) := \mathsf{Tr}(\ulcorner\mathsf{Tr}(\ulcorner \ldots \mathsf{Tr}(\ulcorner\varphi\urcorner)\urcorner)\urcorner)$$

---

16 The idea of analysing semantic paradoxes via equations is already found in Wen (2001). Walicki (2009), Dyrkolbotn & Walicki (2014), Walicki (2017) combined this idea with a graph-theoretical analysis of sentences that employs a pointer structure closely related to the one developed by Gaifman (1988, 1992, 2000). Their approach differs from the one presented here in several respects. For instance, the Walicki-Dyrkolbotn approach gives rise to a noncompositional semantics, while the approach I develop here yields a compositional semantics (see §4.6). The two approaches make also a very different use of equations. I do not compare the two approaches any further in the interest of space.

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

with n nested truth predicates prefixed to φ.[17] McGee sentences are, essentially, infinitary liars (more precisely, ω-liars). μ yields the following graph:

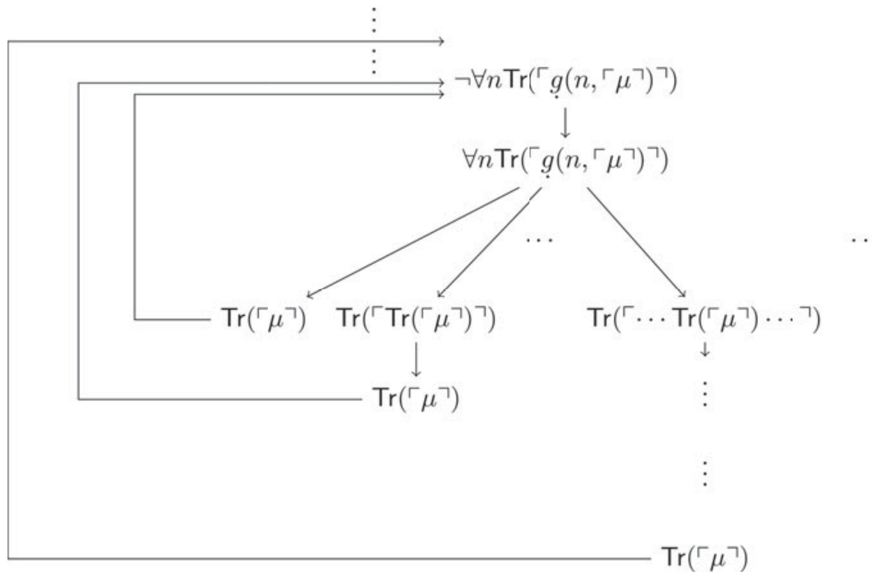

Fig. 5. Graph for the McGee sentence μ.

Once again, only relational information is available. Therefore, one can assign equations to the (infinitely many) sentences appearing in graph 5, according to the evaluation clauses associated with their logical form. The graph yields a single infinitary system, which intuitively works as follows: sentences of the form $\mathsf{Tr}(\ulcorner \cdots \mathsf{Tr}(\ulcorner \mu \urcorner) \cdots \urcorner)$ are required to have all the same value, $\forall n \mathsf{Tr}(\ulcorner g(n, \ulcorner \mu \urcorner) \urcorner)$ is required to be the *infimum* of their values *and* to have the same value of its negation $\neg\forall n \mathsf{Tr}(\ulcorner g(n, \ulcorner \mu \urcorner) \urcorner)$. Again, the only solution of the system is easily seen to be $\mathbf{1/2}$ for every sentence appearing in graph 5.

3.3. Truth-teller–like sentences. In the case of λ, the relational information codified by the equations was turned into a 'standard', numerical value assignment by solving the resulting system. But this is not always possible. Consider the truth-teller sentence τ identical to $\mathsf{Tr}(\ulcorner \tau \urcorner)$. Here is the graph associated with it:

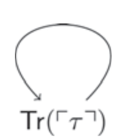

Fig. 6. Graph for the truth-teller sentence $\mathsf{Tr}(\ulcorner \tau \urcorner)$.

As in the liar case, the only semantic information that can be extracted from the graph and the evaluation clause for Tr is relational, i.e., equational. Here is the equation system associated with $\mathsf{Tr}(\ulcorner \tau \urcorner)$:

[17] See McGee (1985). I follow the formulation in Halbach (2011) (p. 157 and following).

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

$$\text{value of } \mathsf{Tr}(\ulcorner\tau\urcorner) = \text{ value of } \mathsf{Tr}(\ulcorner\tau\urcorner).$$

But this system has *more than one solution*. Indeed, every element of any value space is a solution. Therefore, in this case it is not possible to proceed from the relational information determined by graph 6 to an assignment of standard numerical values. In order to encode and express this peculiar feature of τ, i.e., that it can be assigned any 'standard' semantic value available, one can introduce special semantic values. Here, I propose to assign *the equation system itself* to τ, as its semantic value. Equations are quite informative, as far as the semantics of τ is concerned: the system 'expresses' that the only constraints on τ's possible values is that they have to be identical to themselves, and therefore that τ is compatible with any assignment of standard semantic values.

3.4. Revenge sentences. An opposite scenario is given in the case of revenge sentences: their graph determines an equation system with *no solutions* (rather than more than one solution). Consider a language featuring a unary operator D, whose evaluation is governed by the following clause:

$$e(\mathsf{D}(\varphi)) = \mathbf{1} - \min(\mathbf{1}, 2(\mathbf{1} - e(\varphi)))$$

and assume the following numerical value space $V = \{\mathbf{1}, \mathbf{1/2}, \mathbf{0}\}$.[18] Consider a revenge liar sentence λ_{d} identical to $\neg\mathsf{D}(\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner))$, which yields the following graph:

$$\begin{tikzcd}
\neg\mathsf{D}(\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner)) \arrow[d] \\
\mathsf{D}(\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner)) \arrow[d] \\
\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner) \arrow[uu, bend left]
\end{tikzcd}$$

Fig. 7. Graph for the revenge liar sentence $\neg\mathsf{D}(\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner))$.

Here is the associated equation system:[19]

$$\begin{aligned}
x &= \mathbf{1} - y \\
y &= \mathbf{1} - \min(\mathbf{1}, 2(\mathbf{1} - z)) \\
z &= x.
\end{aligned}$$

This system has no solution in $\mathsf{V}_{\text{Ł}} = \{\mathbf{1}, \mathbf{1/2}, \mathbf{0}\}$, although it has a unique solution in a larger numerical value space: $x = \mathbf{2/3} = z$ and $y = \mathbf{1/3}$. Also in this case, no 'standard' value is determined by the equation system: therefore, I assign to revenge sentences their very equation systems as semantic values.

The idea behind assigning equations as semantic values is that they are 'as close as possible' to numerical values. Equations exhibit all the semantic relations determined by truth-teller–like and revenge sentences. In turn, the fact that such semantic relations give rise to equation systems with too many or too few solutions (in the selected value spaces) accounts for the fact that truth-teller–like and revenge sentences admit of too many or

18 This evaluation clause captures the intended semantics of D on the selected value space, i.e., the value of $\mathsf{D}(\varphi)$ is $\mathbf{1}$ if the value of φ is $\mathbf{1}$, and $\mathbf{0}$ otherwise.

19 Where x stands for the value of $\neg\mathsf{D}(\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner))$, y for the value of $\mathsf{D}(\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner))$, and z for the value of $\mathsf{Tr}(\ulcorner\lambda_{\mathsf{d}}\urcorner)$.

too few interpretations via 'standard' semantic values. In conclusion, directly employing equations as semantic values makes it possible to represent in one model the semantic behaviour of truth-teller–like and revenge sentences.[20]

This completes the informal picture of the possible outcomes of the evaluation method I propose here, and the resulting picture matches the taxonomy of sentences outlined in §2.4. To begin with, a sentence is either assigned a numerical semantic value (e.g., $\mathsf{Tr}(\ulcorner t = t \urcorner)$), or an equational value (liar-like, truth-teller–like, and revenge sentences). If equations are assigned, the corresponding systems can have a unique solutions, in which case standard semantic values replace equations (this happens with liar-like sentences). Alternatively, equation systems can either have more than one solution or no solution, in which case equations are kept as semantic values (this happens with truth-teller–like sentences or revenge sentences, respectively). In the next section, I develop a proper semantic theory of truth and paradox, namely an evaluation function (for sufficiently expressive languages) that systematically yields the classification and the semantic value assignments just outlined.

3.5. Intermezzo: loops and nonwell-foundedness. The evaluation procedure outlined here implicitly yields a rather nonstandard answer to the question of whether paradoxical sentences are nonwell-founded. The informal evaluation procedure described so far pictures sentences such as λ, τ and relevantly similar ones as well-founded, in that their decomposition (and the search for information leading to an evaluation) does not lead into an infinite regress. This is because their structure is modelled via *graphs* rather than *trees*, and loops are admissible in the former but not in the latter. Therefore, λ are τ turn out to be well-founded in the more precise sense that their graphs do not have infinitely descending paths. More customarily, λ and τ are decomposed as follows:[21]

$$\begin{array}{ccc}
\neg\mathsf{Tr}(\ulcorner\lambda\urcorner) & \qquad\qquad & \mathsf{Tr}(\ulcorner\tau\urcorner) \\
\downarrow & & \downarrow \\
\mathsf{Tr}(\ulcorner\lambda\urcorner) & & \mathsf{Tr}(\ulcorner\tau\urcorner) \\
\downarrow & & \downarrow \\
\neg\mathsf{Tr}(\ulcorner\lambda\urcorner) & & \mathsf{Tr}(\ulcorner\tau\urcorner) \\
| & & | \\
\vdots & & \vdots
\end{array}$$

In the present approach, λ and τ are decomposed via loops, that 'end', so to speak, the infinitely descending branches of their tree-theoretical representation (see Figures 3 and 6). Nonwell-founded graphs, that is semantic graphs with infinitely descending paths, can be easily obtained if one introduces a *satisfaction* predicate into the language, or suitable recursive functions that make it definable in terms of truth.[22] Then, *Visser–Yablo* sentences become formalisable, and give rise to nonwell-founded graphs. Visser–Yablo sentences ascribe truth, untruth, or some other property to a collection of sentences which in turn ascribe truth, untruth, or some other property to *another* collection of sentences, and so

[20] For further discussion, see §4.7.

[21] See e.g., Yablo (1985, p. 130). For analyses of semantic paradoxes that employ graphs (and therefore feature loops as well), see e.g., Barwise & Etchemendy (1987), Gaifman (1988, 1992, 2000), Schlenker (2007), Walicki (2009), Rabern, Rabern, & Macauley (2013), Cook (2014), Beringer & Schindler (2017).

[22] See e.g., Cook (2014) (p. 22 and ff.).

on, without end. Having a satisfaction predicate (whether primitive or definable) in the language does not alter the evaluation procedure described here.

§4. A unified theory of truth and paradox. The plan of the section is as follows. In §4.1–§4.2, I introduce some graph-theoretical notions and formally define semantic graphs. In §4.3, I fix a semantics for the logical vocabulary, and in §4.4 I provide a semantic construction to assign semantic values to the nodes of semantic graphs. In §4.5, I prove an isomorphism result about semantic graphs, and in §4.6 I use this result to construct the evaluation for truth and paradox that I propose here, called the *canonical evaluation*. In §4.7, I outline some possible variations on the canonical evaluation and some possible further developments.

4.1. Technical preliminaries. Consider a first-order language $\mathcal{L}_{\mathsf{Tr}}$ with identity and a primitive predicate $\mathsf{Tr}(x)$ for 'x is true'. $\mathcal{L}$ is the Tr-free fragment of $\mathcal{L}_{\mathsf{Tr}}$. $\mathcal{L}_{\mathsf{Tr}}$ should be rich enough to encode facts about its own syntax, as in the case of the language of arithmetic. Therefore, I require that $\mathcal{L}_{\mathsf{Tr}}$ satisfy the following requirements:

- It must be possible to define in $\mathcal{L}_{\mathsf{Tr}}$ a coding function $\ulcorner\ \urcorner$ such that for every $\mathcal{L}_{\mathsf{Tr}}$-formula φ, $\ulcorner\varphi\urcorner$ is a closed $\mathcal{L}_{\mathsf{Tr}}$-term. Informally, $\ulcorner\varphi\urcorner$ can be considered as a name of φ.
- For every open $\mathcal{L}_{\mathsf{Tr}}$-formula $\varphi(x)$ there is a term t_φ such that the term $\ulcorner\varphi(t_\varphi/x)\urcorner$ is t_φ, where $\varphi(t_\varphi/x)$ results from replacing every occurrence of x with t_φ in φ.
- $\mathcal{L}_{\mathsf{Tr}}$ has at least one ω-model, i.e., a model isomorphic to the standard model of natural numbers.[23]

The primitive logical constants of $\mathcal{L}_{\mathsf{Tr}}$ are $\neg, \wedge, \rightarrow, \forall$ ($\vee, \leftrightarrow, \exists$ are defined as usual). $\mathsf{CTer}_{\mathcal{L}_{\mathsf{Tr}}}$, $\mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}$, $\mathsf{For}_{\mathcal{L}_{\mathsf{Tr}}}$ indicate the sets of (codes of) closed terms, sentences, and formulae of $\mathcal{L}_{\mathsf{Tr}}$, respectively. Lowercase Latin letters are used as meta-variables for closed terms of $\mathcal{L}_{\mathsf{Tr}}$ (and open terms, if specifically stated). Lowercase Greek letters are used as meta-variables for sentences of $\mathcal{L}_{\mathsf{Tr}}$ (and formulae, if specifically stated). '$\varphi \in \mathcal{L}_{\mathsf{Tr}}$' is a shorthand for '$\varphi \in \mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}$' and 's.t.' is a shorthand for 'such that'. Analogous conventions are in place for the sublanguage $\mathcal{L}$.

I now introduce some basic graph-theoretical notions.[24]

DEFINITION 4.1. *A* directed graph *is a pair* $\langle \mathtt{N}, \mathtt{S}\rangle$, *where* $\mathtt{N}$ *is a nonempty set whose elements are called* nodes, *and* $\mathtt{S}$ *is a set of ordered pairs of nodes, called* edges. $\mathtt{v}$ *and* $\mathtt{w}$, *possibly with indices, range over nodes. For any directed graph* $\langle \mathtt{N}, \mathtt{S}\rangle$, *define the following notions*:

- *A directed graph* $\langle \mathtt{N}^\dagger, \mathtt{S}^\dagger\rangle$ *is a* subgraph *of* $\langle \mathtt{N}, \mathtt{S}\rangle$, *in symbols* $\langle \mathtt{N}^\dagger, \mathtt{S}^\dagger\rangle \subseteq_g \langle \mathtt{N}, \mathtt{S}\rangle$, *if* $\mathtt{N}^\dagger \subseteq \mathtt{N}$ *and* $\mathtt{S}^\dagger \subseteq \mathtt{S}$.
- *A* standard path *is a finite, nonempty tuple of alternating nodes and edges, that begins and ends with nodes, and where every edge connects the two nodes that precede and follow it. More intuitively, it is an object of the following form*:

[23] Since an ω-model is also *acceptable* in the sense of Moschovakis (1974, p. 22), this requirement ensures that it is possible to add a satisfaction predicate to $\mathcal{L}$.

[24] For comprehensive surveys of graph-theoretical notions and results, see Bondy & Murty (2008), Diestel (2010).

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

$$\Big\langle \mathtt{v}_1, \langle \mathtt{v}_1, \mathtt{v}_2\rangle, \mathtt{v}_2, \langle \mathtt{v}_2, \mathtt{v}_3\rangle \ldots, \mathtt{v}_{n-2}, \langle \mathtt{v}_{n-2}, \mathtt{v}_{n-1}\rangle, \mathtt{v}_{n-1}, \langle \mathtt{v}_{n-1}, \mathtt{v}_n\rangle, \mathtt{v}_n\Big\rangle.$$

A simple path*, or just a* path*, is the tuple of edges* $\mathtt{P} \subseteq \mathtt{S}$ *resulting from removing the nodes in a standard path.*[25] *More intuitively, it is an object of the following form*:

$$\Big\langle \langle \mathtt{v}_1, \mathtt{v}_2\rangle, \langle \mathtt{v}_2, \mathtt{v}_3\rangle, \ldots, \langle \mathtt{v}_{n-2}, \mathtt{v}_{n-1}\rangle, \langle \mathtt{v}_{n-1}, \mathtt{v}_n\rangle\Big\rangle.$$

A path $\mathtt{P} \subseteq \mathtt{S}$ *is* from $\mathtt{v}$ to $\mathtt{w}$ *if* $\mathtt{v}$ *is the first element of its first edge and* $\mathtt{w}$ *is the second element of its last edge. A path* $\mathtt{P} \subseteq \mathtt{S}$ *is* maximal *if there is no path* $\mathtt{P}' \subseteq \mathtt{S}$ *s.t.* $\mathtt{P} \neq \mathtt{P}'$ *and* $\mathtt{P} \subseteq \mathtt{P}'$.

- *For a set of edges* $\mathtt{P} \subseteq \mathtt{S}$*, let* $\mathtt{Nodes}(\mathsf{P})$ *denote the set of nodes in* P*. A path* $\mathtt{P} \subseteq \mathtt{S}$ *is a* loop *if, for every* $\mathtt{v} \in \mathtt{Nodes}(\mathtt{P})$*, there is a path* $\mathtt{P}' \subseteq \mathtt{S}$ *from* $\mathtt{v}$ *to* $\mathtt{v}$ *s.t.* $\mathtt{Nodes}(\mathtt{P}') = \mathtt{Nodes}(\mathtt{P})$ *where no node except* $\mathtt{v}$ *occurs twice.*[26]
- *A path* $\mathtt{P}$ *is* straight *if no subpath of* $\mathtt{P}$ *is a loop. For* $\mathtt{P} \subseteq \mathtt{S}$ *a straight path from* $\mathtt{w}$ *to* $\mathtt{v}$*, the set*

$$\mathtt{Pred}_{\mathtt{w}}(\mathtt{v}) := \mathtt{Nodes}(\mathtt{P}\ \textit{minus its last pair})$$

is the set of the predecessors of $\mathtt{v}$ from $\mathtt{w}$. *If* $\mathtt{v}_1 \in \mathtt{Pred}_{\mathtt{w}}(\mathtt{v})$*, then* $\mathtt{v}$ *is a* successor *of* $\mathtt{v}_1$ *from* $\mathtt{w}$*. If* $\langle \mathtt{v}_1, \mathtt{v}\rangle \in \mathtt{P}$*, then* $\mathtt{v}$ *is an* immediate successor *of* $\mathtt{v}_1$*, and* $\mathtt{v}_1$ *is an* immediate predecessor *of* $\mathtt{v}$.[27] $\mathtt{I\text{-}Prec}_{\mathtt{w}}(\mathtt{v})$ *and* $\mathtt{I\text{-}Succ}_{\mathtt{w}}(\mathtt{v})$ *denote, respectively, the immediate predecessor and the immediate successor(s) of* $\mathtt{v}$ *from* $\mathtt{w}$.
- *A node* $\mathtt{v} \in \mathtt{N}$ *is a* dead end *if there is no node* $\mathtt{w} \in \mathtt{N}$ *s.t.* $\langle \mathtt{v}, \mathtt{w}\rangle \in \mathtt{S}$.

4.2. Semantic graphs. For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, I define one directed, labelled, rooted graph $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi\rangle$ and its labelling function $\mathtt{L}_\varphi$, i.e., the function that assign $\mathcal{L}_{\mathsf{Tr}}$-formulae to the nodes of $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi\rangle$. In order to define semantic graphs, I start from the definition of three inductive jumps, respectively corresponding to the operations of extending an arbitrary rooted graph with the results of decomposing sentences whose main operator is unary, binary, or the universal quantifier.

DEFINITION 4.2. *For every directed graph* $\langle \mathtt{N}, \mathtt{S}\rangle$ *labelled with* $\mathcal{L}_{\mathsf{Tr}}$*-sentences and a root note* $\mathtt{r}$*, and for every function* $\mathtt{L} : \mathtt{N}_\varphi \mapsto \mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}$ *(i.e., every labelling function), define the following sets by simultaneous induction*:[28]

(I) $\mathtt{v}_\mathtt{i} \in \mathtt{N}^{\mathsf{U}}$, $\langle \mathtt{v}, \mathtt{v}_\mathtt{i}\rangle \in \mathtt{S}^{\mathsf{U}}$ *and* $\langle \mathtt{v}_\mathtt{i}, \sigma\rangle \in \mathtt{L}^{\mathsf{U}}$ *if*:

(1) $\mathtt{v}_\mathtt{i} \in \mathtt{N}$, $\langle \mathtt{v}, \mathtt{v}_\mathtt{i}\rangle \in \mathtt{S}$ *and* $\langle \mathtt{v}_\mathtt{i}, \sigma\rangle \in \mathtt{L}$; *or*

(2) $\mathtt{v} \in \mathtt{N}$, $\mathtt{L}(\mathtt{v}) = \neg\psi$, *and*

(2.1) *for every* $\mathtt{w} \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}) \neq \psi$*, and* $\sigma = \psi$*, or*

(2.L) *for some* $\mathtt{w} \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}) = \psi$*, and* $\mathtt{v}_\mathtt{i} = \mathtt{w}$; *or*

---

25 I give, and use, the simplified notion of path in order to shorten the proof of some results below, but they could be proven also with the standard definition of path in place.

26 What I call ‘loop’ is more commonly referred to as ‘simple cycle’ in graph theory.

27 In what follows, I only consider *rooted* graphs, with the parameter $\mathtt{w}$ always being the root node.

28 The labels U, B, and I are for a *u*nary, *b*inary, and *i*nfinitary decomposition of sentences respectively.

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

(3) $\mathtt{v} \in \mathtt{N}$, $\mathtt{L}(\mathtt{v}) = \mathsf{Tr}(t)$, *t denotes the code of an* $\mathcal{L}_{\mathsf{Tr}}$*-sentence* ψ*, and*

(3.1) *for every* $\mathtt{w} \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}) \neq \psi$*, and* $\sigma = \psi$*, or*
(3.L) *for some* $\mathtt{w} \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}) = \psi$*, and* $\mathtt{v}_\mathtt{i} = \mathtt{w}$.

(II) $\mathtt{v}_\mathtt{i}, \mathtt{v}_\mathtt{j} \in \mathtt{N}^{\mathsf{B}}$, $\langle \mathtt{v}, \mathtt{v}_\mathtt{i} \rangle, \langle \mathtt{v}, \mathtt{v}_\mathtt{j} \rangle \in \mathtt{S}^{\mathsf{B}}$ *and* $\langle \mathtt{v}_\mathtt{i}, \sigma_0 \rangle, \langle \mathtt{v}_\mathtt{j}, \sigma_1 \rangle \in \mathtt{L}^{\mathsf{B}}$ *if*:

(4) $\mathtt{v}_\mathtt{i}, \mathtt{v}_\mathtt{j} \in \mathtt{N}$, $\langle \mathtt{v}, \mathtt{v}_\mathtt{i} \rangle, \langle \mathtt{v}, \mathtt{v}_\mathtt{j} \rangle \in \mathtt{S}$ *and* $\langle \mathtt{v}_\mathtt{i}, \sigma_0 \rangle, \langle \mathtt{v}_\mathtt{j}, \sigma_1 \rangle \in \mathtt{L}$; *or*
(5) $\mathtt{v} \in \mathtt{N}$, $\mathtt{L}(\mathtt{v}) = \psi \circ \chi$*, and*

(5.1) *for every* $\mathtt{w} \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}) \neq \psi$ *and* $\mathtt{L}(\mathtt{w}) \neq \chi$, $\sigma_0 = \psi$ *and* $\sigma_1 = \chi$*; or*
(5.L.a) *for every* $\mathtt{w} \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}) \neq \psi$*, and for a* $\mathtt{w}_0 \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}_0) = \chi$, $\sigma_0 = \psi$, $\mathtt{w}_0 = \mathtt{v}_\mathtt{j}$*, and* $\sigma_1 = \chi$*; or*
(5.L.b) *for every* $\mathtt{w} \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}) \neq \chi$*, and for a* $\mathtt{w}_0 \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}_0) = \psi$, $\sigma_0 = \chi$, $\mathtt{w}_0 = \mathtt{v}_\mathtt{j}$*, and* $\sigma_1 = \psi$*; or*
(5.L.c) *for a* $\mathtt{w}_0 \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}_0) = \psi$*, for a* $\mathtt{w}_1 \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}_1) = \chi$, $\mathtt{w}_0 = \mathtt{v}_\mathtt{i}$, $\mathtt{w}_1 = \mathtt{v}_\mathtt{j}$, $\sigma_0 = \psi$*, and* $\sigma_1 = \chi$.

(III) $\mathtt{v}_\mathtt{n} \in \mathtt{N}^{\mathsf{I}}$, $\langle \mathtt{v}, \mathtt{v}_\mathtt{n} \rangle \in \mathtt{S}^{\mathsf{I}}$ *and* $\langle \mathtt{v}_\mathtt{n}, \sigma_n \rangle \in \mathtt{L}^{\mathsf{I}}$ *if*:

(6) $\mathtt{v}_\mathtt{n} \in \mathtt{N}$, $\langle \mathtt{v}, \mathtt{v}_\mathtt{n} \rangle \in \mathtt{S}$ *and* $\langle \mathtt{v}_\mathtt{n}, \sigma_n \rangle \in \mathtt{L}$ *if; or*
(7) $\mathtt{v} \in \mathtt{N}$, $\mathtt{L}(\mathtt{v}) = \forall x \chi(x)$*, and for every* $n \in \omega$ *(letting* t_x *be the x-th term in a nonrepeating enumeration of* $\mathsf{CTer}_{\mathcal{L}_{\mathsf{Tr}}}$*)*

(7.1) *for every* $\mathtt{w} \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}) \neq \chi(t_n)$*, and* $\sigma_n = \chi(t_n)$*; or*
(7.L) *for some* $\mathtt{w}_0 \in \mathtt{Pred}_{\mathtt{r}}(\mathtt{v})$, $\mathtt{L}(\mathtt{w}_0) = \chi(t_n)$*, and* $\sigma_n = \chi(t_n)$.

Call '*looping* clauses' the clauses with an 'L' in their label. Definition 4.2 specifies inductively the process of adding edges and labelled nodes to a given graph. For instance, if $\langle \mathtt{N}, \mathtt{S} \rangle$ has a node $\mathtt{v}$ labelled with $\neg\psi$, clause (I)(2) yields a super-graph of that also contains a node $\mathtt{v}_\mathtt{i}$ labelled with ψ and the new edge $\langle \mathtt{v}, \mathtt{v}_\mathtt{i} \rangle$ (if no loop arises), or that contains the new edge $\langle \mathtt{v}, \mathtt{w}_0 \rangle$ (if a loop arises with a predecessor $\mathtt{w}_0$ of $\mathtt{v}$ in $\langle \mathtt{N}, \mathtt{S} \rangle$ that is labelled with ψ).

In order to define semantic graphs, one just needs to put together the clauses of Definition 4.2. For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, the semantic graph $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$ and its labelling function $\mathtt{L}_\varphi$ are the results of applying the clauses of Definition 4.2 to a graph only consisting of a node $\mathtt{r}$ (the root), labelled with φ, until a fixed point is reached.

DEFINITION 4.3. *For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$*, the* semantic graph generated by φ*,* $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$*, and its* labelling function*,* $\mathtt{L}_\varphi : \mathtt{N}_\varphi \mapsto \mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}$*, are the least fixed points of the following simultaneous inductive definition*:

– *At stage* 0*, put*:

$$\mathtt{N}^0_\varphi = \{\mathtt{r}\}; \qquad \mathtt{S}^0_\varphi = \varnothing; \qquad \mathtt{L}^0_\varphi = \{\langle \mathtt{r}, \varphi \rangle\}.$$

– *For an arbitrary successor stage* $\alpha + 1$*, put*:

$$\mathtt{N}^{\alpha+1}_\varphi = (\mathtt{N}^\alpha_\varphi)^{\mathsf{U}} \cup (\mathtt{N}^\alpha_\varphi)^{\mathsf{B}} \cup (\mathtt{N}^\alpha_\varphi)^{\mathsf{I}}; \quad \mathtt{S}^{\alpha+1}_\varphi = (\mathtt{S}^\alpha_\varphi)^{\mathsf{U}} \cup (\mathtt{S}^\alpha_\varphi)^{\mathsf{B}} \cup (\mathtt{S}^\alpha_\varphi)^{\mathsf{I}};$$
$$\mathtt{L}^{\alpha+1}_\varphi = (\mathtt{L}^\alpha_\varphi)^{\mathsf{U}} \cup (\mathtt{L}^\alpha_\varphi)^{\mathsf{B}} \cup (\mathtt{L}^\alpha_\varphi)^{\mathsf{I}}.$$

– *For δ a limit ordinal, put*:

$$\mathtt{N}_{\varphi}^{\delta} = \bigcup_{\alpha<\delta} \mathtt{N}_{\varphi}^{\alpha}; \qquad \mathtt{S}_{\varphi}^{\delta} = \bigcup_{\alpha<\delta} \mathtt{S}_{\varphi}^{\alpha}; \qquad \mathtt{L}_{\varphi}^{\delta} = \bigcup_{\alpha<\delta} \mathtt{L}_{\varphi}^{\alpha}.$$

Finally, put (*where* Ord *is the class of all ordinals*):

$$\mathtt{N}_{\varphi} = \bigcup_{\alpha\in\mathsf{Ord}} \mathtt{N}_{\varphi}^{\alpha}; \qquad \mathtt{S}_{\varphi} = \bigcup_{\alpha\in\mathsf{Ord}} \mathtt{S}_{\varphi}^{\alpha}; \qquad \mathtt{L}_{\varphi} = \bigcup_{\alpha\in\mathsf{Ord}} \mathtt{L}_{\varphi}^{\alpha}.$$

The above definition simply regiments and generalises the informal process that was followed when building semantic graphs in §3.1–§3.4. The graphs that result from Definition 4.3 are exactly of the kind employed in §3.1–§3.4. The next results are immediate from Definitions 4.2–4.3: they ensure that, for every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, the sets $\mathtt{N}_{\varphi}$, $\mathtt{S}_{\varphi}$, and $\mathtt{L}_{\varphi}$ are positive elementary in $\mathtt{N}_{\varphi}^{0}$, $\mathtt{S}_{\varphi}^{0}$, and $\mathtt{L}_{\varphi}^{0}$, and establish their existence and uniqueness.

LEMMA 4.4. *Let* $\mathtt{Pred}_{\varphi}(\mathtt{v})$ *denote the set of predecessors of* $\mathtt{v}$ *from the root note of the semantic graph generated by* φ. *For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ *and every* $\mathtt{v} \in \mathtt{N}_{\varphi}$, *the set* $\mathtt{Pred}_{\varphi}(\mathtt{v})$ *is finite.*

COROLLARY 4.5. *For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ *and* $\mathtt{v} \in \mathtt{N}_{\varphi}$, *there are at most finitely many* $\mathtt{v}_{\mathtt{i}} \in \mathtt{N}_{\varphi}$ *s.t.* $\mathtt{v}_{\mathtt{i}}$ *is a predecessor of* $\mathtt{v}$ *and* $\langle \mathtt{v}, \mathtt{v}_{\mathtt{i}} \rangle \in \mathtt{S}_{\varphi}$.

COROLLARY 4.6. *For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, *there is* exactly one *semantic graph* $\langle \mathtt{N}_{\varphi}, \mathtt{S}_{\varphi} \rangle$ *and* exactly one *labelling function* $\mathtt{L}_{\varphi}$.

4.3. A semantics for the logical vocabulary. The taxonomy of ‘paradoxical’ sentences and the characterisation of their semantic behaviour to be offered here remains structurally unaltered across every compositional semantics for the logical vocabulary of $\mathcal{L}_{\mathsf{Tr}}$. However, in order to give a semantic theory of truth and paradox proper, a semantics for the logical vocabulary must be selected. Therefore, for the sake of presentation, I adopt *Łukasiewicz logics*.[29] This choice is suggested by the fact that revenge paradoxes can be already constructed in theories of naïve truth interpreted via Łukasiewicz semantics, without adding further logical or semantic vocabulary. In fact, Łukasiewicz logic is incompatible with naïve truth, unless revenge paradoxes are blocked in some way.[30] And this is exactly what happens in the proposed semantics, where revenge-theoretical sentences are assigned equations as values, since they cannot be assigned numerical values (see §2.3 and p. 11).

A Łukasiewicz numerical value space $\mathsf{V}_{\text{Ł}}$ is either $\{\mathbf{0}, \mathbf{1/n+1}, \ldots, \mathbf{n/n+1}, \mathbf{1}\}$ (for n an odd positive integer), or the set of reals in the unit interval $[\mathbf{0}, \mathbf{1}]$. I use the boldface letters $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ to range over elements of $\mathsf{V}_{\text{Ł}}$. Here are the Łukasiewicz evaluation clauses:

$$\begin{aligned}
\text{value of } \neg\psi &= \mathbf{1} - \text{ value of } \psi \\
\text{value of } \psi \wedge \chi &= \mathsf{min}[\text{value of } \psi, \text{value of } \chi] \\
\text{value of } \psi \rightarrow \chi &= \mathsf{min}[\mathbf{1}, (\mathbf{1} - \text{ value of } \psi + \text{ value of } \chi)] \\
\text{value of } \forall x\psi(x) &= \mathsf{inf}[\text{value of } \psi(t_n) \mid n \in \omega].
\end{aligned}$$

---

29 For more on Łukasiewicz logics and semantics, see Gottwald (2001).

30 See Restall (1992), Hájek *et al.*, (2000). *Continuum-valued* Łukasiewicz logic is merely ω-inconsistent with naïve truth, but since I am assuming some ω-model of the base language, ω-inconsistency amounts to a proper inconsistency for the models I consider.

The Łukasiewicz clauses for $\neg$, $\wedge$, and $\forall$ generalise the clauses of partial evaluations (see p. 4) to larger value spaces, while $\rightarrow$ can be used to express a comparison between the values of antecedent and consequent.[31] The above clauses remain constant for any choice of $\mathsf{V}_{Ł}$.

I now define a language $\mathcal{L}_{Ł}$ for every numerical value space $\mathsf{V}_{Ł}$, in order to represent the equations definable from the clauses of Łukasiewicz semantics and naïve truth.

DEFINITION 4.7. *The language $\mathcal{L}_{Ł}$ is composed of the following elements*:

- *Countably infinitely many fresh variables* $\mathsf{Var}_{\mathcal{L}_{Ł}} := \{u_{\varphi_1}, \ldots, u_{\varphi_n}, \ldots\}$, *where* φ_n *is the n-th element of a nonrepeating enumeration of* $\mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}$.
- *A set of fresh constants* $\mathsf{Con}_{\mathcal{L}_{Ł}}$ *that contains exactly one term for each element of the numerical value space* $\mathsf{V}_{Ł}$. *I use the same meta-variables to range over both elements of* $\mathsf{V}_{Ł}$ *and elements of* $\mathsf{Con}_{\mathcal{L}_{Ł}}$.[32]
- *Let* $s_1, \ldots, s_i, s_j, s_m, s_n, \ldots$ *be* $\mathcal{L}_{Ł}$*-terms. Then,* $\mathsf{id}(s_m)$ (*the identity function*), $s_m - s_n$, $\mathsf{min}(s_m, s_n)$, $\mathsf{min}[s_i, (s_j - s_m + s_n)]$, $\mathsf{inf}\{s_1, \ldots, s_m, s_n, \ldots\}$ *are* $\mathcal{L}_{Ł}$*-terms* (*and nothing else is*). *Let* h_i *range over the term-forming functions of* $\mathcal{L}_{Ł}$.
- *Let* s_m, s_n *be* $\mathcal{L}_{Ł}$*-terms. Then* $\mathbf{s_m} =_{Ł} \mathbf{s_n}$ *is an atomic formula of* $\mathcal{L}_{Ł}$ (*and nothing else is*). *Denote this set with* $\mathsf{E}_{Ł}$, *and call its element* Łukasiewicz equations. *I use the boldface letter* $\mathbf{e}$, *possibly with indices, to range over elements of* $\mathsf{E}_{Ł}$, *and* $\mathbf{E}$, *possibly with indices, to range over elements of* $\mathcal{P}(\mathsf{E}_{Ł})$ (*i.e., subsets of* $\mathsf{E}_{Ł}$).

Now that I have formally constructed semantic graphs and defined the semantic values, both numerical and equational, to be employed in the semantics, I turn to the assignment of semantic values to nodes in semantic graphs, making the process described in §3 formally precise.

4.4. Evaluations of nodes in semantic graphs. I start off by distinguishing some kinds of nodes in semantic graphs.

DEFINITION 4.8. *For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, *a node* $\mathtt{v}$ *in* $\mathtt{N}_\varphi$ *is*:

- *a* dead end, *if there is no edge* $\langle \mathtt{v}, \mathtt{w} \rangle \in \mathtt{S}_\varphi$ (*no arrow departs from* $\mathtt{v}$). $\mathtt{D\text{-}Ends}_\varphi$ *denotes the set of dead ends of a semantic graph* $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$.
- *a* looping end, *if there are only edges* $\langle \mathtt{v}, \mathtt{w} \rangle \in \mathtt{S}_\varphi$ *where* $\mathtt{w}$ *is a predecessor of* $\mathtt{v}$ (*only looping, upwards arrows depart from* $\mathtt{v}$). $\mathtt{L\text{-}Ends}_\varphi$ *denotes the set of looping ends of a semantic graph* $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$.
- *a* simple point *if it is not an end and it is not in any loop.* $\mathtt{S\text{-}Points}_\varphi$ *denotes the set of simple points of a semantic graph* $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$.
- *a* looping point *if it is not an end, it is in a loop, and its immediate predecessor is in a loop.* $\mathtt{L\text{-}Points}_\varphi$ *denotes the set of looping points of a semantic graph* $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$.

31 For the use and relevance of the Łukasiewicz conditional in theories of truth, see Rossi (2016).

32 $\mathsf{Con}_{\mathcal{L}_{Ł}}$ in effect contains the elements of $\mathsf{V}_{Ł}$ to be used as *constants* in solving equation systems definable in $\mathcal{L}_{Ł}$. I do not put $\mathsf{Con}_{\mathcal{L}_{Ł}} = \mathsf{V}_{Ł}$ to avoid confusion in the definition of $\mathcal{L}_{Ł}$, although the two sets are identified in practice. Notice that, if $\mathsf{V}_{Ł} = [\mathbf{0}, \mathbf{1}]$, then $\mathsf{Con}_{\mathcal{L}_{Ł}}$ is uncountable, and this makes the language $\mathcal{L}_{Ł}$ itself uncountable. For simplicity, however, I adopt a countable notation for $\mathcal{L}_{Ł}$, since only countably many different values are assigned in the semantics to be developed, as there are only countably many $\mathcal{L}_{\mathsf{Tr}}$-sentences.

– *a* loop top *it is in a loop and it is the root or its immediate predecessor is a simple point.* $\mathtt{L\text{-}Tops}_\varphi$ *denotes the set of loop tops of a semantic graph* $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$*. If* $\mathtt{v}$ *is in a loop,* $\mathtt{w}$ *is* the loop top of $\mathtt{v}$ *if* $\mathtt{w}$ *is the only loop top in* $\mathtt{Pred}_\varphi(\mathtt{v})$ *and there is a path* $\mathtt{P}$ *from* $\mathtt{w}$ *to* $\mathtt{v}$ *such that* $\mathtt{w}$ *is the only loop top in* $\mathtt{P}$.

LEMMA 4.9. *For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$*, every looping end or looping point in* $\mathtt{N}_\varphi$ *has exactly one loop top in* $\mathtt{N}_\varphi$*.*

Using the above classification, I now show how to assign values to nodes in semantic graphs. Values are assigned in a revision procedure, that replaces (whenever possible) equational values with numerical values. The assignment of equational and numerical values, in turn, is given by two inductive definitions. More specifically, the first inductive construction assigns single equational values, while the second one assigns numerical values (whenever possible) and sets of equations.[33]

I begin with the first inductive construction. It is a 'top-bot' construction, that is it starts assigning an equational value to the root of a graph, and then moves on to assign equations to its successors. Equations are assigned according to the logical form of each node.

DEFINITION 4.10. *Let* $\mathcal{M}$ *be an* ω*-model of* $\mathcal{L}$*. For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ *and* $A \subseteq (\mathtt{N}_\varphi \times \mathsf{E}_{\text{Ł}})$*,* $\langle x, y \rangle \in A^+_\varphi$ *if*:

$x = \mathtt{r}_\varphi$ *or* $x \in \mathtt{N}_\varphi$ *and there is a* $\mathtt{w} \in \mathtt{I\text{-}Prec}_\varphi(x)$ *and an* $\mathbf{e} \in \mathsf{E}_{\text{Ł}}$ *s.t.* $\langle \mathtt{w}, \mathbf{e} \rangle \in A$ *and*:

1. $\mathtt{L}_\varphi(x) = P(t_1, \ldots, t_n)$*,* $P(t_1, \ldots, t_n)$ *is an atomic* $\mathcal{L}$*-sentence,* $\mathcal{M} \models P(t_1, \ldots, t_n)$*, and* $y = (\mathbf{u}_\psi {=_{\text{Ł}}} \mathbf{1})$*; or*
2. $\mathtt{L}_\varphi(x) = P(t_1, \ldots, t_n)$*,* $P(t_1, \ldots, t_n)$ *is an atomic* $\mathcal{L}$*-sentence and* $\mathcal{M} \not\models P(t_1, \ldots, t_n)$*, or* $\mathtt{L}_\varphi(x) = \mathsf{Tr}(t)$ *and* t *does not denote the code of a* $\mathcal{L}_{\mathsf{Tr}}$*-sentence in* $\mathcal{M}$*, and* $y = (\mathbf{u}_\psi {=_{\text{Ł}}} \mathbf{0})$*; or*
3. $\mathtt{L}_\varphi(x) = \neg\psi$*, and* $y = (\mathbf{u}_{\neg\psi} {=_{\text{Ł}}} \mathbf{1} - \mathbf{u}_\psi)$*; or*
4. $\mathtt{L}_\varphi(x) = \psi \wedge \chi$*, and* $y = (\mathbf{u}_{\psi\wedge\chi} {=_{\text{Ł}}} \mathsf{min}(\mathbf{u}_\varphi, \mathbf{u}_\psi))$*; or*
5. $\mathtt{L}_\varphi(x) = \psi \to \chi$*, and* $y = (\mathbf{u}_{\psi\to\chi} {=_{\text{Ł}}} \mathsf{min}[1, (1 - \mathbf{u}_\varphi + \mathbf{u}_\psi)])$*; or*
6. $\mathtt{L}_\varphi(x) = \forall x \chi(x)$*, and* $y = (\mathbf{u}_{\forall x\chi(x)} {=_{\text{Ł}}} \mathsf{inf}\{\mathbf{u}_{t_k} \mid k \in \omega\})$*; or*
7. $\mathtt{L}_\varphi(x) = \mathsf{Tr}(\ulcorner \psi \urcorner)$*, and* $y = (\mathbf{u}_{\mathsf{Tr}(\ulcorner\psi\urcorner)} {=_{\text{Ł}}} \mathbf{u}_\psi)$*.*

LEMMA 4.11. *For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ *and every* $A \subseteq (\mathtt{N}_\varphi \times \mathsf{E}_{\text{Ł}})$*, the definition of* A^+_φ *is* positive elementary in *the following sets*: $\mathsf{E}_{\text{Ł}}$, $\langle \mathtt{S}_\varphi, \mathtt{N}_\varphi \rangle$, $\mathtt{L}_\varphi$, $\mathsf{Var}_{\mathcal{L}_{\text{Ł}}}$, $\mathsf{Con}_{\mathcal{L}_{\text{Ł}}}$, $\{\mathtt{Im\text{-}Prec}_\varphi(\mathtt{v}) \mid \mathtt{v} \in \mathtt{N}_\varphi\}$.

I now re-write Definition 4.10 via an operator on subsets of $(\mathtt{N}_\varphi \times \mathsf{E}_{\text{Ł}})$, which will be useful later.

DEFINITION 4.12. *For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$*, let* $\overline{Q}_\varphi$ *indicate the sets in which the definition of* A^+_φ *is positive elementary, as per Lemma 4.11. For* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ *and* $S \subseteq (\mathtt{N}_\varphi \times \mathsf{E}_{\text{Ł}})$*, let* $\zeta_\varphi(x, y, S, \overline{Q}_\varphi)$ *be the right-hand side of Definition* 4.10*. Let* $\Phi_\varphi : \mathcal{P}(\mathtt{N}_\varphi \times \mathsf{E}_{\text{Ł}}) \longmapsto \mathcal{P}(\mathtt{N}_\varphi \times \mathsf{E}_{\text{Ł}})$ *be the operator defined as*:

$$\Phi_\varphi(S) := \{\langle x, y \rangle \in \mathtt{N}_\varphi \times \mathsf{E}_{\text{Ł}} \mid \zeta_\varphi(x, y, S, \overline{Q}_\varphi)\}.$$

[33] I am grateful to Emmanuel Chemla, whose observations suggested to employ two distinct inductive constructions, simplifying the previous version of the semantics.

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

Put, for $S \subseteq \mathtt{N}_\varphi \times \mathsf{E}_{\text{Ł}}$ *and* γ *limit*:

$$\Phi_\varphi^{\alpha+1}(S) := \Phi(\Phi_\varphi^\alpha(S)), \qquad \Phi_\varphi^\gamma(S) := \bigcup_{\alpha<\gamma} \Phi_\varphi^\alpha(S).$$

LEMMA 4.13. *For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$*, the operator* Φ_φ *is monotone.*

Let $\mathsf{I}_{\Phi_\varphi}$ denote the smallest fixed point of Φ_φ, that is:

$$\mathsf{I}_{\Phi_\varphi} := \bigcup_{\alpha \in \mathsf{Ord}} \Phi_\varphi^\alpha(\varnothing).$$

I now turn to the second inductive construction, which incorporates $\mathsf{I}_{\Phi_\varphi}$.

DEFINITION 4.14. *Let* $\mathcal{M}$ *be an* ω*-model of* $\mathcal{L}$*. For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ *and* $B \subseteq \mathtt{N}_\varphi \times (\mathsf{V}_{\text{Ł}} \cup \mathcal{P}(\mathsf{E}_{\text{Ł}}))$*,* $\langle x, y\rangle \in B_\varphi^*$ *if*:

1. $\langle x, y\rangle \in B$; *or*
2. $x \in \mathtt{N}_\varphi$, $\mathtt{L}_\varphi(x) = P(t_1, \ldots, t_n)$, $P(t_1, \ldots, t_n)$ *is an atomic* $\mathcal{L}$*-sentence,* $\mathcal{M} \models P(t_1, \ldots, t_n)$*, and* $y = \mathbf{1}$; *or*
3. $x \in \mathtt{N}_\varphi$, $\mathtt{L}_\varphi(x) = P(t_1, \ldots, t_n)$, $P(t_1, \ldots, t_n)$ *is an atomic* $\mathcal{L}$*-sentence and* $\mathcal{M} \not\models P(t_1, \ldots, t_n)$*, or* $\mathtt{L}_\varphi(x) = \mathsf{Tr}(t)$ *and* t *does not denote the code of a* $\mathcal{L}_{\mathsf{Tr}}$*-sentence in* $\mathcal{M}$*, and* $y = \mathbf{0}$; *or*
4. $x \in \mathtt{N}_\varphi$, $\mathtt{L}_\varphi(x) = \psi \wedge \chi$ *or* $\mathtt{L}_\varphi(x) = \forall x\chi(x)$ *and there is a* $\mathtt{v_m} \in \mathtt{I\text{-}Succ}_\varphi(x)$ *s.t.* $\langle \mathtt{v_m}, \mathbf{0}\rangle \in B$*, and* $y = \mathbf{0}$; *or*
5. $x \in \mathtt{N}_\varphi$, $\mathtt{L}_\varphi(x) = \psi \rightarrow \chi$ *and there is a* $\mathtt{v_m} \in \mathtt{I\text{-}Succ}_\varphi(x)$ *s.t.* $\mathtt{L}_\varphi(\mathtt{v_m}) = \psi$ *and* $\langle \mathtt{v_m}, \mathbf{0}\rangle \in B$*, or there is a* $\mathtt{v_n} \in \mathtt{I\text{-}Succ}_\varphi(x)$ *s.t.* $\mathtt{L}_\varphi(\mathtt{v_n}) = \chi$ *and* $\langle \mathtt{v_n}, \mathbf{1}\rangle \in B$*, and* $y = \mathbf{1}$; *or*
6. $x \in \mathtt{N}_\varphi$, $x \in \mathtt{S\text{-}Points}_\varphi$, $\mathtt{L}_\varphi(x) = \neg\psi$*, and there is a* $\mathtt{v_m} \in \mathtt{I\text{-}Succ}_\varphi(x)$*, and a* $\mathbf{j} \in \mathsf{V}_{\text{Ł}}$ *s.t.* $\mathtt{L}_\varphi(\mathtt{v_m}) = \psi$, $\langle \mathtt{v_m}, \mathbf{j}\rangle \in B$*, and* $y = \mathbf{1} - \mathbf{j}$; *or*
7. $x \in \mathtt{N}_\varphi$, $x \in \mathtt{S\text{-}Points}_\varphi$, $\mathtt{L}_\varphi(x) = \psi \wedge \chi$*, and there are* $\mathtt{v_m}, \mathtt{v_n} \in \mathtt{I\text{-}Succ}_\varphi(x)$ *and* $\mathbf{j}, \mathbf{k} \in \mathsf{V}_{\text{Ł}}$ *s.t.* $\langle \mathtt{v_m}, \mathbf{j}\rangle \in B$ *and* $\langle \mathtt{v_n}, \mathbf{k}\rangle \in B$*, and* $y = \min(\mathbf{j}, \mathbf{k})$; *or*
8. $x \in \mathtt{N}_\varphi$, $x \in \mathtt{S\text{-}Points}_\varphi$, $\mathtt{L}_\varphi(x) = \psi \rightarrow \chi$*, and there are* $\mathtt{v_m}, \mathtt{v_n} \in \mathtt{I\text{-}Succ}_\varphi(x)$ *s.t.* $\mathtt{L}_\varphi(\mathtt{v_m}) = \psi$ *and* $\mathtt{L}_\varphi(\mathtt{v_n}) = \chi$*, and there are* $\mathbf{j}, \mathbf{k} \in \mathsf{V}_{\text{Ł}}$ *s.t.* $\langle \mathtt{v_m}, \mathbf{j}\rangle \in B$ *and* $\langle \mathtt{v_n}, \mathbf{k}\rangle \in B$*, and* $y = \min[\mathbf{1}, (\mathbf{1} - \mathbf{j} + \mathbf{k})]$; *or*
9. $x \in \mathtt{N}_\varphi$, $x \in \mathtt{S\text{-}Points}_\varphi$, $\mathtt{L}_\varphi(x) = \forall x\chi(x)$*, and for every* $m \in \omega$ *there is a* $\mathtt{v_m} \in \mathtt{I\text{-}Succ}_\varphi(x)$ *and a* $\mathbf{i_m} \in \mathsf{V}_{\text{Ł}}$ *s.t.* $\mathtt{L}_\varphi(\mathtt{v_m}) = \chi(t_m)$, $\langle \mathtt{v_m}, \mathbf{i_m}\rangle \in B$*, and* $y = \inf\{\mathbf{i_m} \in \mathsf{V}_{\text{Ł}} \mid \langle \mathtt{v_m}, \mathbf{i_m}\rangle \in B$ *and* $\mathtt{v_m} \in \mathtt{I\text{-}Succ}_\varphi(x)\}$; *or*
10. $x \in \mathtt{N}_\varphi$, $x \in \mathtt{S\text{-}Points}_\varphi$, $\mathtt{L}_\varphi(x) = \mathsf{Tr}(\ulcorner\psi\urcorner)$*, and there is a* $\mathtt{v_m} \in \mathtt{I\text{-}Succ}_\varphi(x)$ *and a* $\mathbf{j} \in \mathsf{V}_{\text{Ł}}$ *s.t.* $\langle \mathtt{v_m}, \mathbf{j}\rangle \in B$*, and* $y = \mathbf{j}$; *or*
11. $x \in \mathtt{N}_\varphi$, $x \in \mathtt{L\text{-}Tops}_\varphi$*, and* y *is the set of* $\mathbf{e} \in \mathsf{E}_{\text{Ł}}$ *s.t.* $\langle \mathtt{w}, \mathbf{e}\rangle \in \mathsf{I}_{\Phi_\varphi}$ *and either* $\mathtt{w} = x$ *or* x *is the loop top of* $\mathtt{w}$; *or*
12. $x \in \mathtt{N}_\varphi$, $x \in \mathtt{L\text{-}Points}_\varphi$ *or* $x \in \mathtt{L\text{-}Ends}_\varphi$*, there is a* $\mathbf{E} \in \mathcal{P}(\mathsf{E}_{\text{Ł}})$ *s.t.* $\langle \mathtt{v}, \mathbf{E}\rangle \in B$ *and* $\mathtt{v}$ *is the loop top of* x*, and* y *is the set of* $\mathbf{e} \in \mathsf{E}_{\text{Ł}}$ *s.t.* $\langle \mathtt{w}, \mathbf{e}\rangle \in \mathsf{I}_{\Phi_\varphi}$ *and* $\mathtt{w}$ *and* x *have the same loop top*; *or*
13. $x \in \mathtt{N}_\varphi$, $x \in \mathtt{S\text{-}Points}_\varphi$ *and there is a* $\mathtt{v} \in \mathtt{I\text{-}Succ}_\varphi(x)$ *and* $\mathbf{E} \in \mathcal{P}(\mathsf{E}_{\text{Ł}})$ *s.t.* $\langle \mathtt{v}, \mathbf{E}\rangle \in B$*, and* y *is the union of the set of* $\mathbf{e} \in \mathsf{E}_{\text{Ł}}$ *s.t.* $\langle x, \mathbf{e}\rangle \in \mathsf{I}_{\Phi_\varphi}$ *and the set of* $\mathbf{e} \in \mathsf{E}_{\text{Ł}}$ *s.t.* $\langle \mathtt{v}, \mathbf{e}\rangle \in \mathsf{I}_{\Phi_\varphi}$; *or*

14. *$x \in \mathtt{N}_\varphi$, $x \in \mathtt{L\text{-}Tops}_\varphi$, the equation system given by the set* $\{\mathbf{e} \in \mathsf{E}_{Ł} \mid \langle \mathtt{w}, \mathbf{e}\rangle \in \mathsf{I}_{\Phi_\varphi}$ *and either* $\mathtt{w} = x$ *or* x *is the loop top of* $\mathtt{w}\}$ *has a unique solution in* $\mathsf{V}_{Ł}$, *and* y *is the solution for* x *in* $\mathsf{V}_{Ł}$; *or*
15. *$x \in \mathtt{N}_\varphi$, $x \in \mathtt{L\text{-}Points}_\varphi$ or $x \in \mathtt{L\text{-}Ends}_\varphi$, there is a* $\mathbf{k} \in \mathsf{V}_{Ł}$ *s.t.* $\langle \mathtt{v}, \mathbf{k}\rangle \in B$ *and* $\mathtt{v}$ *is the loop top of* x, *the equation system given by the set* $\{\mathbf{e} \in \mathsf{E}_{Ł} \mid \langle \mathtt{w}, \mathbf{e}\rangle \in \mathsf{I}_{\Phi_\varphi}$ *and* $\mathtt{w}$ *and* x *have the same loop top*$\}$ *has a unique solution in* $\mathsf{V}_{Ł}$, *and* y *is the solution for* x *in* $\mathsf{V}_{Ł}$; *or*

Definition 4.14 deals with all the possible cases in which a numerical value or an equation system is assigned to a node, and is therefore somewhat intricate. However, it merely regiments the heuristics described in §3, and its working is actually quite simple. First, dead ends are assigned either value **1** or **0**, according to the selected ω-model of the base language (items 2 and 3). Second, conjunctions with a **0**-valued conjunct, universal quantifications with a **0**-valued instance, and conditionals with a **0**-valued antecedent or a **1**-valued consequent are assigned a numerical value (items 4 and 5). Third, simple points whose immediate successors are assigned a numerical value are also assigned a numerical value, compositionally (items 6–10). Fourth, equation systems are assigned starting from loop tops, employing the equations assigned in the smallest fixed point of Φ_φ, i.e., $\mathsf{I}_{\Phi_\varphi}$ (items 11–13). Finally, equation systems are solved and numerical values are assigned whenever possible (items 14 and 15).

LEMMA 4.15. *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ and $B \subseteq \mathtt{N}_\varphi \times (\mathsf{V}_{Ł} \cup \mathcal{P}(\mathsf{E}_{Ł}))$, the definition of B^*_φ is* positive elementary in *the following sets*: $\mathsf{V}_{Ł}$, $\mathsf{E}_{Ł}$, $\langle \mathtt{S}_\varphi, \mathtt{N}_\varphi\rangle$, $\mathtt{L}_\varphi$, $\mathsf{Var}_{\mathcal{L}_{Ł}}$, $\mathsf{Con}_{\mathcal{L}_{Ł}}$, $\mathtt{L\text{-}Ends}_\varphi$, $\mathtt{S\text{-}Points}_\varphi$, $\mathtt{L\text{-}Points}_\varphi$, $\mathtt{L\text{-}Tops}_\varphi$, $\{\mathtt{I\text{-}Succ}_\varphi(\mathtt{v}) \mid \mathtt{v} \in \mathtt{N}_\varphi\}$, $\{\mathtt{Pred}_\varphi(\mathtt{v}) \mid \mathtt{v} \in \mathtt{N}_\varphi\}$, $\{\langle \mathtt{v}, \mathtt{N}^\dagger\rangle \in \mathtt{N}_\varphi \times \mathcal{P}(\mathtt{N}_\varphi) \mid \mathtt{v}$ *is the loop top of the nodes in* $\mathtt{N}^\dagger\}$, $\{\mathsf{E} \in \mathcal{P}(\mathsf{E}_{Ł}) \mid \mathsf{E}$ *has a unique solution in* $\mathsf{V}_{Ł}\}$.

I re-write also Definition 4.14 via an operator on subsets of $\mathtt{N}_\varphi \times (\mathsf{V}_{Ł} \cup \mathcal{P}(\mathsf{E}_{Ł}))$, to be used later.

DEFINITION 4.16. *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, let $\overline{R}_\varphi$ indicate the sets in which the definition of B^*_φ is positive elementary, as per Lemma 4.15. For $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ and $S \subseteq \mathtt{N}_\varphi \times (\mathsf{V}_{Ł} \cup \mathcal{P}(\mathsf{E}_{Ł}))$, let $\vartheta_\varphi(x, y, S, \overline{R}_\varphi)$ be the right-hand side of Definition* 4.14. *Let $\Psi_\varphi : \mathcal{P}(\mathtt{N}_\varphi \times (\mathsf{V}_{Ł} \cup \mathcal{P}(\mathsf{E}_{Ł}))) \longmapsto \mathcal{P}(\mathtt{N}_\varphi \times (\mathsf{V}_{Ł} \cup \mathcal{P}(\mathsf{E}_{Ł})))$ be the operator defined as*:

$$\Psi_\varphi(S) := \{\langle x, y\rangle \in \mathtt{N}_\varphi \times (\mathsf{V}_{Ł} \cup \mathcal{P}(\mathsf{E}_{Ł})) \mid \vartheta_\varphi(x, y, S, \overline{R}_\varphi)\}.$$

LEMMA 4.17. *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, the operator Ψ_φ is monotone and increasing.*

A fundamental idea of the heuristics outlined in §3 is that nodes are *first* assigned equation systems, and *then* numerical values (when equation systems are solved, if they have a unique solution). However, this amounts to revising a previously assigned value, and cannot be done via an inductive construction such as the one given in Definition 4.14. Inductive constructions can only *add* numbers and equation systems: in order to replace the latter with the former, a *revision construction* is required. This is provided by the next definition.

DEFINITION 4.18. *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ and ordinal δ, let e^δ_φ be defined as follows* (*for γ limit*):

$$
\begin{aligned}
e_\varphi^0 &:= \mathsf{I}_{\Phi_\varphi}\\
e_\varphi^{\alpha+1} &:= \Psi_\varphi(e_\varphi^\alpha) \setminus \{\langle x, y\rangle \in \mathrm{N}_\varphi \times \mathcal{P}(\mathsf{E}_{Ł}) \mid \\
&\qquad\qquad \langle x, y\rangle \in \Psi_\varphi(e_\varphi^\alpha) \textit{ and for some } \mathbf{k} \in \mathsf{V}_{Ł},\ \langle x, \mathbf{k}\rangle \in \Psi_\varphi(e_\varphi^\alpha)\}\\
e_\varphi^\gamma &:= \{\langle x, y\rangle \in \mathrm{N}_\varphi \times (\mathsf{V}_{Ł} \cup \mathcal{P}(\mathsf{E}_{Ł})) \mid \textit{there is an } \alpha < \gamma \textit{ s.t.}\\
&\qquad\qquad \textit{for all } \alpha \leq \beta < \gamma,\ \langle x, y\rangle \in e_\varphi^\beta\}.
\end{aligned}
$$

The key element of the above revision sequence is the successor case. In short, whenever $\Psi_\varphi(e_\varphi^\alpha)$ assigns *both* an equation system and a numerical value to a node in N_φ, $e_\varphi^{\alpha+1}$ removes the equation system and keeps only the numerical value. As it turns out, this suffices to ensure that every set e_φ^α is a *function*, and that equational values are revised and replaced with numerical values as outlined in the heuristics of §3. The limit case then ensures that, as ordinals grow, the functions e_φ^α converge to a limit, that is they have fixed points. These facts are formulated more precisely and collected in the following proposition.

PROPOSITION 4.19.

(I) *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ and every ordinal α:*

1. *There is exactly one e_φ^α.*
2. *e_φ^α is a* function*, i.e., for every $\mathrm{v} \in \mathrm{N}_\varphi$ and every $\mathbf{v_0}, \mathbf{v_1} \in \mathsf{V}_{Ł} \cup \mathcal{P}(\mathsf{E}_{Ł})$:*

   $$\textit{if } \langle \mathrm{v}, \mathbf{v_0}\rangle \in e_\varphi^\alpha \textit{ and } \langle \mathrm{v}, \mathbf{v_1}\rangle \in e_\varphi^\alpha, \textit{ then } \mathbf{v_0} = \mathbf{v_1}.$$

   I write '$e_\varphi^\alpha(\mathrm{v}) = \mathbf{v}$' if $\langle \mathrm{v}, \mathbf{v}\rangle \in e_\varphi^\alpha$, and '$e_\varphi^\alpha = e_\varphi^\beta$' if, for every $\mathrm{v} \in \mathrm{N}_\varphi$, $e_\varphi^\alpha(\mathrm{v}) = e_\varphi^\beta(\mathrm{v})$.
3. *For every $\mathrm{v} \in \mathrm{N}_\varphi$, if $e_\varphi^\alpha(\mathrm{v}) = \mathbf{k}$ for a $\mathbf{k} \in \mathsf{V}_{Ł}$, then for every $\beta > \alpha$, $e_\varphi^\beta(\mathrm{v}) = \mathbf{k}$.*
4. *For every $\mathrm{v} \in \mathrm{N}_\varphi$, if $e_\varphi^\alpha(\mathrm{v}) = \mathbf{E}$ for $\mathbf{E} \in \mathcal{P}(\mathsf{E}_{Ł})$, then for every $\beta > \alpha$, if there is no $\mathbf{k} \in \mathsf{V}_{Ł}$ s.t. $e_\varphi^\beta(\mathrm{v}) = \mathbf{k}$, then $e_\varphi^\beta(\mathrm{v}) = \mathbf{E}$.*

(II) *There is* exactly one *ordinal δ_0 s.t. for every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, $e_\varphi^{\delta_0}$ is a* fixed point *of the functions e_φ^α, i.e., for every $\delta \geq \delta_0$ and $\varphi \in \mathcal{L}_{\mathsf{Tr}}$:*

$$e_\varphi^{\delta_0} = e_\varphi^\delta$$

I indicate $e_\varphi^{\delta_0}$ simply as e_φ.

4.5. Loop-isomorphisms. We have seen how semantic graphs are precisely constructed, and how their nodes are assigned a semantic value, formalising the picture outlined in §3. However, as I argued in §1, in order to provide a full account of semantic paradoxes, one needs an *evaluation function* for $\mathcal{L}_{\mathsf{Tr}}$-*sentences*. But so far we only have many evaluation functions for *labelled nodes*—one such function per graph. Therefore, the evaluations defined on nodes have to be turned into a single evaluation defined on sentences.

To see this, consider the nodes in the graph generated by λ (see Figure 3): they are assigned value $\mathbf{1/2}$ by the evaluation function associated with that very graph, i.e., e_λ (see §3.2). However, e_λ does not tell us anything about the value assigned to a node labelled with λ occurring in *another* semantic graph, i.e., in the graph generated by another sentence. But if one thinks that the *sentence* λ should be assigned value $\mathbf{1/2}$ and adopts a compositional semantics, presumably one also thinks that the *sentence* $\lambda \wedge \neg(s = s)$ should be assigned value $\mathbf{0}$. But e_λ does not give us this information, because it is not a function that evaluates $\mathcal{L}_{\mathsf{Tr}}$-sentences—it evaluates only the nodes of $\mathtt{N}_\lambda$.

In order to 'weave together' the evaluation functions defined on nodes $(e_{\varphi_1}, e_{\varphi_2}, \ldots)$ and construct a single evaluation defined on sentences, I show that the functions e_φ have the following *robustness* property: all the nodes with the same label are assigned the same value by their respective evaluation functions (Proposition 4.23). This makes it possible to define a *canonical evaluation* for sentences (Definition 4.24), by taking the value of φ to be the value of a uniformly chosen node labelled with φ – for simplicity, I will take the root node of $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$. In order to prove the robustness property, I show that whenever two nodes $\mathtt{v}$ and $\mathtt{w}$ are labelled with the same sentence, they generate subgraphs of their respective semantic graphs $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$ and $\langle \mathtt{N}_\psi, \mathtt{S}_\psi \rangle$ that are structurally similar (Proposition 4.22). Such similarity is used to ensure that the two evaluations e_φ and e_ψ yield the same result on $\mathtt{v}$ and $\mathtt{w}$. The relevant notion of structural similarity is provided by a suitable notion of graph-theoretic isomorphism (Definition 4.21).

To begin with, note that in order to evaluate a node $\mathtt{v}$ of $\mathtt{N}_\varphi$ via the function e_φ, possibly not all of $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$ is relevant. Consider for instance the semantic graph of $\lambda \wedge \neg(s = s)$:

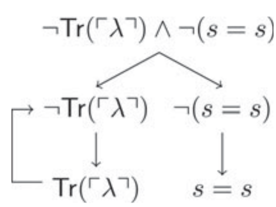

Fig. 8. Graph for $\neg\mathsf{Tr}(\ulcorner\lambda\urcorner) \wedge \neg(s = s)$.

In order to evaluate some of the nodes of graph 8, not all other nodes need to be evaluated: for instance, in order to evaluate the node labelled with $\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)$, the value of the node labelled with $s = s$ is not required. More generally, an inspection of Definitions 4.3 and 4.14 shows that, in order to assign a value to a node $\mathtt{v}$ in $\mathtt{N}_\varphi$, only the nodes that can be *reached from* $\mathtt{v}$ in following the edges, i.e., the arrows, in $\mathtt{S}_\varphi$ are employed in the construction of $e_\varphi(\mathtt{v})$. The next definition makes the notion of reachable nodes more precise.

DEFINITION 4.20 (Reachable nodes and subgraphs). *For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ *and* $\mathtt{v} \in \mathtt{N}_\varphi$, *the set* $\mathtt{R}_\varphi(\mathtt{v})$ *of* nodes reachable from $\mathtt{v}$ within $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$ *is defined thus*:

$$\mathtt{R}_\varphi(\mathtt{v}) := \Big\{ \Big\langle \langle \mathtt{v}, \mathtt{w} \rangle, \ldots, \langle \mathtt{v}', \mathtt{w}' \rangle \Big\rangle \in \bigcup_{n \in \omega} (\mathtt{N}_\varphi \times \mathtt{N}_\varphi)^n \mid \textit{there is a path from } \mathtt{v} \textit{ to } \mathtt{w}' \textit{ in } \langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle \Big\}.$$

Let $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ *and* $\mathtt{v} \in \mathtt{N}_\varphi$. *The* subgraph of $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$ reachable from $\mathtt{v}$ *is the graph* $\langle \mathtt{N}^{\mathtt{v}}_\varphi, \mathtt{S}^{\mathtt{v}}_\varphi \rangle$ *s.t.*:

$$\mathtt{N}^{\mathtt{v}}_\varphi := \textit{the nodes in } \mathtt{R}_\varphi(\mathtt{v}); \qquad \mathtt{S}^{\mathtt{v}}_\varphi := \textit{the edges in } \mathtt{R}_\varphi(\mathtt{v}).$$

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

I now define the required notion of isomorphism between semantic graphs.

DEFINITION 4.21 (Loop-isomorphism). *Let* $\varphi, \psi \in \mathcal{L}_{\mathsf{Tr}}$, $\langle \mathtt{N}^0_\varphi, \mathtt{S}^0_\varphi \rangle \subseteq_g \langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$ *and* $\langle \mathtt{N}^0_\psi, \mathtt{S}^0_\psi \rangle \subseteq_g \langle \mathtt{N}_\psi, \mathtt{S}_\psi \rangle$. $\langle \mathtt{N}^0_\varphi, \mathtt{S}^0_\varphi \rangle$ *and* $\langle \mathtt{N}^0_\psi, \mathtt{S}^0_\psi \rangle$ *are* loop-isomorphic, *in symbols* $\langle \mathtt{N}^0_\varphi, \mathtt{S}^0_\varphi \rangle \cong_l \langle \mathtt{N}^0_\psi, \mathtt{S}^0_\psi \rangle$, *if*:

(i) *for every dead end* (*simple point*) $\mathtt{v} \in \mathtt{N}^0_\varphi$ *there is a dead end* (*simple point*) $\mathtt{w} \in \mathtt{N}^0_\psi$ *s.t.* $\mathtt{L}_\varphi(\mathtt{v}) = \mathtt{L}_\psi(\mathtt{w})$, *and* vice versa *for dead ends and simple points of* $\mathtt{N}^0_\varphi$, *and*

(ii) *for every loop* $\mathtt{P}_1 \subseteq \mathtt{S}^0_\varphi$ *there is a loop* $\mathtt{P}_2 \subseteq \mathtt{S}^0_\psi$ *s.t. for every pair* $\langle \mathtt{v}, \mathtt{v}' \rangle \in \mathtt{P}_2$ *there is a pair* $\langle \mathtt{w}, \mathtt{w}' \rangle \in \mathtt{P}_2$ *s.t.* $\mathtt{L}_\varphi(\mathtt{v}) = \mathtt{L}_\psi(\mathtt{w})$ *and* $\mathtt{L}_\varphi(\mathtt{v}') = \mathtt{L}_\psi(\mathtt{w}')$, *and* vice versa *for loops in* $\mathtt{S}^0_\psi$.

Informally, two (sub)graphs are loop-isomorphic when every dead end (simple node) of one graph is bijectively mapped to a dead end (simple point) of the other graph, preserving the identity of labels (item (i)), and every loop of one graph is bijectively mapped to a loop of the other graph, preserving adjacency and identity of labels (item (ii)). It follows from Definition 4.21 that if $\langle \mathtt{N}^0_\varphi, \mathtt{S}^0_\varphi \rangle$ and $\langle \mathtt{N}^0_\psi, \mathtt{S}^0_\psi \rangle$ are loop-isomorphic, then paths of $\mathtt{S}^0_\varphi$ that only contain simple points and dead ends are bijectively mapped to paths of $\mathtt{S}^0_\psi$ that only contain simple points and dead ends, and that have the same labels in the same order, while loops of $\mathtt{S}^0_\varphi$ are bijectively mapped to loops of $\mathtt{S}^0_\psi$ that have the same labels with the same adjacencies, but that are *possibly rotated*.[34]

I now state the main fact about the loop-isomorphisms of semantic graphs, that is that nodes with identical labels yield loop-isomorphic reachable subgraphs. In other words, a node labelled with φ generates a subgraph of the graph it belongs to that is structurally similar (i.e., loop-isomorphic) to the subgraph generated by any other node labelled with φ, in any other graph.

PROPOSITION 4.22. *For every* $\varphi, \psi \in \mathcal{L}_{\mathsf{Tr}}$, $\mathtt{v} \in \mathtt{N}_\varphi$, $\mathtt{w} \in \mathtt{N}_\psi$:

$$\text{if } \mathtt{L}_\varphi(\mathtt{v}) = \mathtt{L}_\psi(\mathtt{w}), \text{ then } \langle \mathtt{N}^{\mathtt{v}}_\varphi, \mathtt{S}^{\mathtt{v}}_\varphi \rangle \cong_l \langle \mathtt{N}^{\mathtt{w}}_\psi, \mathtt{S}^{\mathtt{w}}_\psi \rangle.$$

This result, in turn, makes it possible to prove that the evaluations defined on nodes are robust in the sense described above, that is in the sense that they assign identical values to nodes with identical labels.

PROPOSITION 4.23. *For every* $\varphi, \psi \in \mathcal{L}_{\mathsf{Tr}}$, $\mathtt{v} \in \mathtt{N}_\varphi$, $\mathtt{w} \in \mathtt{N}_\psi$, *and* $\mathbf{v} \in \mathsf{V}_{\mathsf{Ł}} \cup \mathcal{P}(\mathsf{E}_{\mathsf{Ł}})$, *if* $\mathtt{L}_\varphi(\mathtt{v}) = \mathtt{L}_\psi(\mathtt{w})$:

1. *there is an* α *s.t.* $e^\alpha_\varphi(\mathtt{v}) = \mathbf{v}$ *if and only if there is a* β *s.t.* $e^\beta_\psi(\mathtt{w}) = \mathbf{v}$, *and*
2. $e_\varphi(\mathtt{v}) = e_\psi(\mathtt{w})$.

4.6. The canonical evaluation. Proposition 4.23 makes it possible to speak of the value of a *sentence* φ, rather than the value of a node labelled with φ in some semantic graph. More precisely, the value of a sentence φ can be taken to be the value that any

[34] Propositions 4.22 and 4.23 could have been proven employing a more standard notion of isomorphism between labelled directed graphs, i.e., the existence of a bijection preserving adjacency of nodes and identity of labels (and only identity of labels in the case of a graph with empty edges), but this would have made the proofs longer.

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

evaluation e_ψ whose domain includes a node labelled with φ assigns to any such node—by Proposition 4.23, all such nodes receive the same value. So, I define a *canonical evaluation* that takes the value of φ to be the value of a canonically selected node labelled with φ – for simplicity, I take it to be the value of the root node of $\langle \mathtt{N}_\varphi, \mathtt{S}_\varphi \rangle$, the graph generated by φ.

DEFINITION 4.24. *The* canonical evaluation *is the function* $\mathscr{C} : \mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}} \longmapsto \mathsf{V}_{Ł} \cup \mathcal{P}(\mathsf{E}_{Ł})$ *defined as*:

$$\mathscr{C}(\varphi) := e_\varphi(\mathtt{r}).$$

The canonical evaluation $\mathscr{C}$ is the semantic theory of truth and paradox I propose in this article. In this section, I review its main properties.

4.6.1. The canonical evaluation as a theory of paradox. To begin with, $\mathscr{C}$ obeys the Łukasiewicz clauses for introducing and eliminating logical constants.

PROPOSITION 4.25. *For every* $\varphi, \psi \in \mathcal{L}_{\mathsf{Tr}}$ *and* $\chi(x) \in \mathsf{For}_{\mathcal{L}_{\mathsf{Tr}}}$, *the following holds* ($\Longleftrightarrow$ *stands for the meta-linguistic 'if and only if', and* $\Longrightarrow$ *for 'if . . . then'*) :

$$\begin{aligned}
\mathscr{C}(\neg\varphi) = \mathbf{1} &\iff \mathscr{C}(\varphi) = \mathbf{0}\\
\mathscr{C}(\varphi \wedge \psi) = \mathbf{1} &\iff \mathscr{C}(\varphi) = \mathbf{1} \textit{ and } \mathscr{C}(\psi) = \mathbf{1}\\
\mathscr{C}(\varphi \to \psi) = \mathbf{1} &\iff \mathscr{C}(\varphi) = \mathbf{0},\\
&\qquad \textit{or } \mathscr{C}(\psi) = \mathbf{1},\\
&\qquad \textit{or } \mathscr{C}(\varphi) = \mathbf{j}, \mathscr{C}(\psi) = \mathbf{k}, \textit{and } \mathbf{j} \leq \mathbf{k}\\
\mathscr{C}(\forall x \chi(x)) = \mathbf{1} &\iff \mathscr{C}(\chi(t_k)) = \mathbf{1} \textit{ for all } t_k \in \mathsf{CTer}_{\mathcal{L}_{\mathsf{Tr}}}\\
\mathscr{C}(\mathsf{Tr}(\ulcorner \varphi \urcorner)) = \mathbf{1} &\iff \mathscr{C}(\varphi) = \mathbf{1}.
\end{aligned}$$

In addition, modus ponens *holds for the canonical evaluation*:

$$\mathscr{C}(\varphi) = \mathbf{1} \textit{ and } \mathscr{C}(\varphi \to \psi) = \mathbf{1} \implies \mathscr{C}(\psi) = \mathbf{1}.$$

The next result generalises Proposition 4.25 to the whole value space $\mathsf{V}_{Ł} \cup \mathcal{P}(\mathsf{E}_{Ł})$, providing a full picture of how the canonical evaluation interprets $\mathcal{L}_{\mathsf{Tr}}$-sentences.

PROPOSITION 4.26. *Let* $\mathcal{M}$ *be an* ω*-model of* $\mathcal{L}$. *For* $A \subseteq \mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}$, *let* $\mathbf{E}(A)$ *be the set* $\{\mathbf{e} \in \mathsf{E}_{Ł} \mid \mathbf{e} \in \mathscr{C}(\varphi),$ *for* $\varphi \in A\}$. *For all* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, *the following hold*:

$$\mathscr{C}(s = t) = \begin{cases} \mathbf{1}, & \textit{if } \mathcal{M} \models s = t,\\ \mathbf{0}, & \textit{if } \mathcal{M} \not\models s = t \end{cases}$$

$$\mathscr{C}(\neg\varphi) = \begin{cases} \mathbf{1} - \mathscr{C}(\varphi), & \textit{if } \mathscr{C}(\varphi) \in \mathsf{V}_{Ł},\\ \{\mathbf{u}_{\neg\varphi} =_{Ł} \mathbf{1} - \mathbf{s}_\varphi\} \cup \mathbf{E}(\{\varphi\}), & \textit{if } \mathscr{C}(\varphi) \in \mathcal{P}(\mathsf{E}_{Ł}) \end{cases}$$

$$\mathscr{C}(\varphi \wedge \psi) = \begin{cases} \mathbf{0}, \ \textit{if } \mathscr{C}(\varphi) = \mathbf{0} \textit{ or } \mathscr{C}(\psi) = \mathbf{0}\\ \min(\mathscr{C}(\varphi), \mathscr{C}(\psi)), \ \textit{if } \mathscr{C}(\varphi), \mathscr{C}(\psi) \in \mathsf{V}_{Ł},\\ \{\mathbf{u}_{\varphi\wedge\psi} =_{Ł} \min(\mathbf{s}_\varphi, \mathbf{s}_\psi)\} \cup \mathbf{E}(\{\varphi, \psi\}),\\ \qquad \textit{if } \begin{cases} \mathscr{C}(\varphi), \mathscr{C}(\psi) \in \mathcal{P}(\mathsf{E}_{Ł}), \textit{ or}\\ \mathscr{C}(\varphi) \in \mathcal{P}(\mathsf{E}_{Ł}), \mathscr{C}(\psi) \in \mathsf{V}_{Ł} \setminus \{\mathbf{0}\}, \textit{ or}\\ \mathscr{C}(\psi) \in \mathcal{P}(\mathsf{E}_{Ł}), \mathscr{C}(\varphi) \in \mathsf{V}_{Ł} \setminus \{\mathbf{0}\} \end{cases} \end{cases}$$

$$\mathscr{C}(\varphi \to \psi) = \begin{cases} \mathbf{1}, \ \textit{if}\ \mathscr{C}(\varphi) = \mathbf{0}\ \textit{or}\ \mathscr{C}(\psi) = \mathbf{1} \\ \min[\mathbf{1}, (\mathbf{1} - \mathscr{C}(\varphi) + \mathscr{C}(\psi))], \ \textit{if}\ \mathscr{C}(\varphi), \mathscr{C}(\psi) \in \mathsf{V}_{\mathsf{Ł}}, \\ \{\mathbf{u}_{\varphi \to \psi} =_{\mathsf{Ł}} \mathsf{min}[\mathbf{1}, (\mathbf{1} - \mathbf{s}_\varphi + \mathbf{s}_\psi)]\} \cup \mathbf{E}(\{\varphi, \psi\}), \\ \qquad \textit{if} \begin{cases} \mathscr{C}(\varphi), \mathscr{C}(\psi) \in \mathcal{P}(\mathsf{E}_{\mathsf{Ł}}), \textit{ or} \\ \mathscr{C}(\varphi) \in \mathcal{P}(\mathsf{E}_{\mathsf{Ł}}), \mathscr{C}(\psi) \in \mathsf{V}_{\mathsf{Ł}} \setminus \{\mathbf{1}\}, \textit{ or} \\ \mathscr{C}(\psi) \in \mathcal{P}(\mathsf{E}_{\mathsf{Ł}}), \mathscr{C}(\varphi) \in \mathsf{V}_{\mathsf{Ł}} \setminus \{\mathbf{0}\} \end{cases} \end{cases}$$

$$\mathscr{C}(\forall x \chi(x)) = \begin{cases} \mathbf{0}, \ \textit{if}\ \mathscr{C}(\chi(t_k)) = \mathbf{0}\ \textit{for a}\ k \in \omega \\ \inf\{\mathscr{C}(\chi(t_k)) \mid k \in \omega\}, \ \textit{if for all}\ k \in \omega, \mathscr{C}(\chi(t_k)) \in \mathsf{V}_{\mathsf{Ł}}, \\ \{\mathsf{u}_{\forall x \chi(x)} =_{\mathsf{Ł}} \mathsf{inf}\{\mathbf{s}_{\chi(t_k)} | k \in \omega\}\} \cup \mathbf{E}(\{\chi(t_n) \mid n \in \omega\}), \\ \qquad \textit{if for all}\ k \in \omega, \mathscr{C}(\chi(t_k)) \in \mathsf{V}_{\mathsf{Ł}} \cup \mathcal{P}(\mathsf{E}_{\mathsf{Ł}}) \setminus \{\mathbf{0}\}, \\ \qquad \textit{and for some}\ n \in \omega, \mathscr{C}(\chi(t_n)) \in \mathcal{P}(\mathsf{E}_{\mathsf{Ł}}) \end{cases}$$

$$\mathscr{C}(\mathsf{Tr}(\ulcorner \varphi \urcorner)) = \begin{cases} \mathscr{C}(\varphi), \ \textit{if}\ \mathscr{C}(\varphi) \in \mathsf{V}_{\mathsf{Ł}}, \\ \{\mathbf{u}_{\mathsf{Tr}(\ulcorner \varphi \urcorner)} =_{\mathsf{Ł}} \mathbf{s}_\varphi\} \cup \mathbf{E}(\{\varphi\}), \ \textit{if}\ \mathscr{C}(\varphi) \in \mathcal{P}(\mathsf{E}_{\mathsf{Ł}}). \end{cases}$$

The above result summarises the possible outcomes of applications of $\mathscr{C}$ to a sentence φ: either φ is assigned a numerical value or an equation system. The clauses for equation systems may appear strange at first, but they merely formalise the heuristics described in §3. For example, the negation clause $\mathscr{C}(\neg\varphi) = \{\mathbf{u}_{\neg\varphi} =_{\mathsf{Ł}} \mathbf{1} - \mathbf{u}_\varphi\} \cup \mathbf{E}(\{\varphi\})$ tells us that if the immediate subcomponent of $\neg\varphi$ (that is φ) is assigned the equation system $\mathbf{E}(\{\varphi\})$, then φ is assigned the equation system that results from adding an equation corresponding to the logical form of $\neg\varphi$ to $\mathbf{E}(\{\varphi\})$. Similarly, if at least one of $\mathscr{C}(\varphi)$ and $\mathscr{C}(\psi)$ is an equation system, and none of $\mathscr{C}(\varphi)$ and $\mathscr{C}(\psi)$ is $\mathbf{0}$, then $\mathscr{C}(\varphi \wedge \psi)$ is an equation system adding to whatever equation systems are associated with φ and ψ (whether just one of φ and ψ is associated with an equation systems, or both of them are) the equation expressing that the value of $\varphi \wedge \psi$ is the minimum of the values of φ and ψ. And so on.

This does not necessarily mean that the equation system assigned to φ has more equations than the system assigned to its subcomponents. In fact, the subcomponents of φ may be assigned the same equation system as φ, as it is to be expected e.g., in the case of paradoxical sentences generating loops. Consider for example the value of $\mathscr{C}(\neg\mathsf{Tr}(\ulcorner\lambda\urcorner))$ in a *classical* value space $\mathsf{V}_{\mathsf{Ł}} = \{\mathbf{1}, \mathbf{0}\}$:

$$\begin{aligned} \mathscr{C}(\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)) &= \{\mathbf{u}_{\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)} =_{\mathsf{Ł}} \mathbf{1} - \mathbf{u}_{\mathsf{Tr}(\ulcorner\lambda\urcorner)}\} \cup \mathbf{E}(\{\mathsf{Tr}(\ulcorner\lambda\urcorner)\}) \\ &= \{\mathbf{u}_{\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)} =_{\mathsf{Ł}} \mathbf{1} - \mathbf{u}_{\mathsf{Tr}(\ulcorner\lambda\urcorner)}\} \cup \\ &\qquad \{\mathbf{u}_{\mathsf{Tr}(\ulcorner\lambda\urcorner)} =_{\mathsf{Ł}} \mathbf{u}_{\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)}, \mathbf{u}_{\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)} =_{\mathsf{Ł}} \mathbf{1} - \mathbf{u}_{\mathsf{Tr}(\ulcorner\lambda\urcorner)}\} \\ &= \{\mathbf{u}_{\mathsf{Tr}(\ulcorner\lambda\urcorner)} =_{\mathsf{Ł}} \mathbf{u}_{\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)}, \mathbf{u}_{\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)} =_{\mathsf{Ł}} \mathbf{1} - \mathbf{u}_{\mathsf{Tr}(\ulcorner\lambda\urcorner)}\} \\ &= \mathscr{C}(\mathsf{Tr}(\ulcorner\lambda\urcorner)). \end{aligned}$$

The equation system associated with $\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)$ does not have more equations than the system associated with $\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)$. This is the expected result: both λ and $\neg\lambda$ should be associated with the same value, that is the same equation system, which is unsolvable in $\mathsf{V}_{\mathsf{Ł}} = \{\mathbf{1}, \mathbf{0}\}$.

It is now clear that the canonical evaluation yields the classification of $\mathcal{L}_{\mathsf{Tr}}$-sentences described in §2 (see especially §2.4) and §3. Classical numerical values are assigned to nonparadoxical sentences (§3.1). Nonclassical numerical values are assigned to liar-like sentences (§2.1, §3.2). Equation systems are assigned to sentences that are compatible with too many numerical values (truth-teller–like sentences, §2.2, §3.3) or too few numerical values (revenge sentences §2.3, §3.4).

Finally, notice that the canonical evaluation is *compositional*: the value of $\mathscr{C}(\varphi)$ (whether numerical or equational) depends on the values that $\mathscr{C}$ assigns to the immediate subcomponents of φ (where, by NAÏVETÉ, ψ is considered to be a subcomponent of $\mathsf{Tr}(\ulcorner\psi\urcorner)$).

4.6.2. The canonical evaluation as a theory of truth. The next results show that the canonical evaluation also constitutes a semantic theory of naïve truth. To begin with, $\mathscr{C}$ validates the INTERSUBSTITUTIVITY of truth and the T-SCHEMA for every sentence receiving a numerical value (see §2, p. 3).

LEMMA 4.27. *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, if φ^{Tr} is the result of substituting (possibly nonuniformly) a subformula ψ of φ with $\mathsf{Tr}(\ulcorner\psi\urcorner)$ or* vice versa*, then (for $\mathbf{k} \in \mathsf{V}_{Ł}$):*

$$\mathscr{C}(\varphi) = \mathbf{k} = \mathscr{C}(\varphi^{\mathsf{Tr}}).$$

LEMMA 4.28. *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, there is a $\mathbf{k} \in \mathsf{V}_{Ł}$ s.t.:*

$$\mathscr{C}(\varphi) = \mathbf{k} \text{ if and only if } \mathscr{C}(\varphi \leftrightarrow \mathsf{Tr}(\ulcorner\varphi\urcorner)) = \mathbf{1}.$$

Moreover, $\mathscr{C}$ includes some arguably good candidates to determine the extension of a naïve truth predicate, such as the smallest fixed point of Kripke (1975)'s theory (strong Kleene version).

PROPOSITION 4.29. *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$:*

- *if φ is in the extension of Tr in the least Kripkean fixed point for $\mathcal{L}_{\mathsf{Tr}}$, then $\mathscr{C}(\varphi) = \mathbf{1}$, and*
- *if φ is in the anti-extension of Tr in the least Kripkean fixed point for $\mathcal{L}_{\mathsf{Tr}}$, then $\mathscr{C}(\varphi) = (\mathbf{0})$.*

Clearly, the converse of the claims in Proposition 4.29 does not hold.

4.6.3. The canonical evaluation, determinateness, and revenge. The canonical evaluation recovers a natural partial version of every Łukasiewicz semantics, generalised with the inclusion of equational values. However, every finite-valued Łukasiewicz semantics is known to be inconsistent with naïve truth, while continuum-valued Łukasiewicz semantics is inconsistent with naïve truth over ω-models of the base language.[35] The canonical evaluation avoids these difficulties as follows: the sentences that cannot be consistently assigned a value according to the Łukasiewicz semantics clauses in conjunction with naïve truth, i.e., *revenge sentences* for this semantics, are simply assigned an equation system which is unsolvable in the selected value space.

For instance, the following *iterated Curry sentences* (where $\perp$ is some false sentence) produce a revenge paradox for every finitely-valued Łukasiewicz semantics plus naïve truth:[36]

$$\kappa_0 := \mathsf{Tr}(\ulcorner\kappa_0\urcorner) \to \perp$$
$$\kappa_{j+1} := \mathsf{Tr}(\ulcorner\kappa_{j+1}\urcorner) \to \kappa_j.$$

[35] See Restall (1992), Hájek *et al.,* (2000).

[36] For more details, see Field (2008, chap. 4).

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

However, the canonical evaluation assigns them numerical values *if possible* (that is, whenever the numerical space is sufficiently large), and unsolvable equation systems otherwise. At a glance:

$$\mathscr{C}(\kappa_0) = \begin{cases} \mathbf{E} \in \mathcal{P}(\mathsf{E}_{\text{Ł}}), \text{ if } \mathsf{V}_{\text{Ł}} = \{\mathbf{0}, \mathbf{1}\} \\ \mathbf{1/2}, \text{ if } \mathsf{V}_{\text{Ł}} = \{\mathbf{0}, \mathbf{1/n+1}, \ldots, \mathbf{n/n+1}, \mathbf{1}\} \text{ for all odd } n \\ \qquad \text{or } \mathsf{V}_{\text{Ł}} = [\mathbf{0}, \mathbf{1}] \end{cases}$$

$$\mathscr{C}(\kappa_{j+1}) = \begin{cases} \mathbf{E} \in \mathcal{P}(\mathsf{E}_{\text{Ł}}), \text{ if } \mathsf{V}_{\text{Ł}} \subseteq \{\mathbf{0}, \mathbf{1/k+1}, \ldots, \mathbf{k/k+1}, \mathbf{1}\} \text{ for } k < 2^{j+1} \\ \mathbf{2^{j+1}-1/2^{j+1}}, \text{ if } \mathsf{V}_{\text{Ł}} = \{\mathbf{0}, \mathbf{1/k+1}, \ldots, \mathbf{k/k+1}, \mathbf{1}\} \text{ for } k \geq 2^{j+1}, \\ \qquad \text{or } \mathsf{V}_{\text{Ł}} = [\mathbf{0}, \mathbf{1}]. \end{cases}$$

In a similar way, the canonical evaluation blocks the revenge sentences employed to prove the next result, assigning them an unsolvable equation system (see the proof in the Appendix).

PROPOSITION 4.30 (Restall (1992)). *There is no continuum-valued Łukasiewicsz evaluation for $\mathcal{L}_{\mathsf{Tr}}$ that*: (*i*) *agrees with an ω-model for $\mathcal{L}$, and* (*ii*) *validates the* T-SCHEMA *or* INTERSUBSTITUTIVITY.

Restall's omega-inconsistency can be proven via a bivalent notion of determinateness, which is definable in Łukasiewicz semantics.[37] Since the canonical evaluation 'blocks' the applications of bivalent determinateness that are prone to yield revenge sentences (by assigning them equational values), it can consistently feature a *partial* bivalent determinateness operator, i.e., an operator that works bivalently on every sentence receiving a numerical value.

DEFINITION 4.31. *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, put*: $\mathsf{D}(\varphi) := \neg(\varphi \to \neg\varphi)$. *Let $\mathsf{D}^n(\varphi)$ be a string of n iterations of* D *applied to φ. Let* Nt *be a univalent recursive ordinal notation system, whose range is* $\mathsf{Ord}_{\mathsf{Nt}}$. *Let r be the primitive recursive function from positive integers and sentences to sentences s.t. $r(n, \ulcorner\varphi\urcorner) = \mathsf{D}^n(\varphi)$, and put $\mathsf{D}^{\omega}(\varphi) := \forall n \mathsf{Tr}\ulcorner \underset{\cdot}{r}(n, \ulcorner\varphi\urcorner)\urcorner$. Define a determinateness hierarchy* à la *Field for ordinals in* $\mathsf{Ord}_{\mathsf{Nt}}$.[38]

PROPOSITION 4.32. *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ and every $\mathsf{V}_{\text{Ł}}$, if $\mathscr{C}(\varphi) = \mathbf{k}$, for $\mathbf{k} \in \mathsf{V}_{\text{Ł}}$, then*:

1. *For all ordinals $\alpha \in \mathsf{Ord}_{\mathsf{Nt}}$, $\mathscr{C}(\mathsf{D}^{\alpha}(\varphi)) \in \mathsf{V}_{\text{Ł}}$. In particular* (*for γ limit*):

$$\mathscr{C}(\mathsf{D}^{\alpha+1}(\varphi)) = \mathbf{1} - \min[\mathbf{1}, (\mathbf{1} - \mathscr{C}(\mathsf{D}^{\alpha}(\varphi)) + \mathbf{1} - \mathscr{C}(\mathsf{D}^{\alpha}(\varphi)))]$$
$$\mathscr{C}(\mathsf{D}^{\gamma}(\varphi)) = \inf\{\mathscr{C}(\mathsf{D}^{\alpha}(\varphi)) | \alpha < \gamma\}$$

2. *There is a unique ordinal $\delta' \in \mathsf{Ord}_{\mathsf{Nt}}$ s.t. for all $\delta \in \mathsf{Ord}_{\mathsf{Nt}}$ greater than or equal to δ'*:

$$\mathscr{C}(\mathsf{D}^{\delta}(\varphi)) = \begin{cases} \mathbf{1} & \textit{if and only if } \mathscr{C}(\varphi) = \mathbf{1} \\ \mathbf{0} & \textit{if and only if } \mathscr{C}(\varphi) \in \mathsf{V}_{\text{Ł}} \textit{ and } \mathscr{C}(\varphi) < \mathbf{1} \end{cases}$$

---

[37] To my knowledge, the determinateness operator described here was introduced, in the context of Łukasiewicz logic, by Field (2008) (pp. 89–92, but see also Field (2003), p. 157 and ff.).

[38] For ordinal notation systems, see Rogers (1987, secs. 11.7–11.8). For determinateness hierarchies, see Field (2008, chap. 22–23). Since Nt is recursive and univalent, the definable iterations of D turn out to be *much* shorter than those in Field (2008), but neither longer iterations nor stronger notation systems are needed, as the proof of Proposition 4.32 shows.

This result shows that, for every $V_Ł$, the canonical evaluation expresses a *unique*, *bivalent* determinateness operator $D^{\delta'}$ that declares that every sentence that has a numerical value other than 'classical truth' (i.e., **1**) is not determinate.[39] There is no 'fuzzy' hierarchy of stronger and stronger determinateness operators (unlike in Field (2008)): they converge to a bivalent operator at a small ordinal (dependent on the cardinality of $V_Ł$).[40] The canonical evaluation is thus immune from a criticism that was advanced against the theory in Field (2008), namely that it recovers a unique notion of truth but splits the notion of determinate truth into a highly unmanageable hierarchy.[41] The canonical evaluation provides one notion of truth and one notion of determinate truth.

4.7. Modifications, extensions, and prospects for future work. The construction that gives rise to the canonical evaluation is rather flexible, and can be subject to several modifications and extensions. Here I informally outline some of them.

- The Łukasiewicz evaluation clauses (see §4.3) can be replaced with different clauses (the value space should also be modified accordingly). A variant of the canonical evaluation can thus be obtained for every *compositional* semantics for the logical vocabulary. However, there is no obvious way to adapt the canonical evaluation to noncompositional semantics, e.g., supervaluations.
- An immediate modification is given by just considering the classical numerical value space, i.e., $\{\mathbf{1}, \mathbf{0}\}$. When restricted to $\{\mathbf{1}, \mathbf{0}\}$, the Łukasiewicz evaluation clauses (as many other nonclassical evaluation clauses) just reduce to the classical ones. In this value space, liar-like and revenge sentences are completely identified. When there are only classical numerical values, liar-like sentences yield equation systems with no solutions, just like revenge sentences. In a slogan: liar-like sentences are revenge sentences for classical semantics. More generally, in any given semantics, the difference between liar-like sentences, i.e., the 'standard' paradoxical sentences, and revenge sentences is that there are enough numerical values to evaluate sentences of the former kind, but of not the latter kind. But this is no 'deep' difference: the same sentence can be classified as 'liar-like' or 'revenge' depending solely on the available values.[42]
- The canonical evaluation displays a 'strong Kleene-style' approach to partiality: that some subformula of φ has a numerical value is in some cases sufficient for φ to have a numerical value. In particular: a conjunction with a **0**-valued conjunct has itself value **0**, and similarly for **0**-valued universally quantified sentences, or **1**-valued conditionals. One could easily modify (indeed: simplify) the construction of the canonical evaluation to give it a 'weak Kleene-style' approach, where a conjunction $\psi \wedge \chi$ in which ψ has value **0** but χ has an equational value has

[39] The uniqueness of $D^{\delta'}$ holds *modulo* the choice of Nt, but the ordinals involved are so small (ω at most, if $V_Ł = [\mathbf{0}, \mathbf{1}]$) that there is virtually no dependence on the specific notation adopted. For details, see the proof in the Appendix.

[40] This treatment of determinateness also avoids the 'trivial collapse' of Field (2008): at no level of ill-behaved iterations of D the resulting operator sends every sentence to **0**.

[41] For this line of criticism, see e.g., Horsten (2012, sec. 10.2).

[42] This also holds for nonnumerical space values, e.g., the value space adopted in Field (2008, chap. 17). The approaches to revenge developed in Cook (2007, 2009), Schlenker (2010), Cook & Tourville (2016) also suggest that the difference between liar-like and revenge paradoxes depends on the available values. See also footnote 13.

itself an equational value (and similarly for conditionals and universally quantified sentences). These two variants—strong and weak Kleene—can, again, be adapted to all compositional evaluation clauses for the logical vocabulary.

– Another immediate extension of the canonical evaluation would consist in treating naïve predicates for satisfaction, denotation, and other semantic notions. In order to extend the canonical evaluation in this way, it would be sufficient to add naïve evaluation clauses for the corresponding semantic predicates. Extending the canonical evaluation to a language with a naïvely interpreted satisfaction predicate would enable one to treat other paradoxical constructions, such as the Visser-Yablo paradoxes.[43]
– Truth-teller–like sentences and revenge sentences are both assigned equation systems, with the former having systems with more than one solution, and the latter having systems with no solutions (in the selected numerical value space). One might want to make the difference between the two cases more explicit in the evaluation itself, assigning them different kinds of values. One could therefore replace the sets of equations employed in Definitions 4.14 and 4.18 with two new conventional values, to be assigned to truth-teller–like and revenge sentences respectively. A form of compositionality is also preserved in this variant. However, while this solution might seem to be more explicit, it strikes me as less informative, less uniform, and less elegant.
– Two liar sentences with different codes (e.g., with different Gödel numbering, if $\ulcorner \cdot \urcorner$ is defined via Gödelization) are assigned, strictly speaking, two different equation systems, even though they only differ for the choice of Ł-variables. But arguably, it might be objected, any two such sentences should be assigned the *very same* equation system.[44] This problem is easily solved, however: it is sufficient to map all sentences which have equation systems that are identical *modulo* renaming of free Ł-variables to some fixed equation system, employing some fixed Ł-variables.
– Propositions 4.25 and 4.26 show that the canonical evaluation can be associated with some notions of consequence, including both logical and truth-theoretical inferences. For instance, say that φ is a $\mathscr{C}_1$-consequence of a set of sentences Γ if, if for every $\psi_i \in \Gamma$, $\mathscr{C}(\psi_i) = \mathbf{1}$, then $\mathscr{C}(\varphi) = \mathbf{1}$. $\mathscr{C}_1$-consequence can be shown to extend strong Kleene logic with a strong rule of conditional-introduction (as per Proposition 4.25).[45] Other choices for the definition of consequence, such as preservation of values greater than or equal to $\mathbf{1/2}$, preservation of the ordering of

[43] See Visser (1989), Yablo (1985, 1993, 2006), and also Priest (1997), Beall (2001), Leitgeb (2002), Bueno & Colyvan (2003), Ketland (2004, 2005), Cook (2006, 2014), Eldridge-Smith (2015), Halbach & Zhang (2017). In Visser-Yablo cases, one has an unending sequence of sentences, each one to the effect that the sentences that come after it are true, untrue, or else. If one thinks that Visser-Yablo cases should be separated from liar-like, truth-teller–like, and revenge sentences, then the canonical evaluation could be modified in order to categorise them differently, distinguishing between paradoxical sentences involving a straightforward circularity or self-reference, and paradoxical sentences involving a form of ungroundedness or nonwell-foundedness.

[44] Thanks to Joel Hamkins and Richard Kimberly Heck for pointing out this potential problem to me.

[45] For strong Kleene logic, see Urquhart (2001). For extensions of strong Kleene logic with conditionals obeying stronger introduction rules, see Rossi (2016).

numerical values, and so on, give rise to generalisations of other logics, such as LP, ST, TS and more.[46] However, all the standard choices for defining consequence only involve numerical values, thus leaving the equational values 'unused'. This has some possibly unexpected consequences. For instance, consider $\mathscr{C}_1$-consequence: since liar-like and revenge sentence don't have value **1**, one might expect that any sentence is a $\mathscr{C}_1$-consequence of a liar-like or a revenge sentence. However, every sentence also follows from truth-teller–like sentences, since such sentences also don't have value **1**. But this might seem counterintuitive. Why should liar-like, revenge, and truth-teller–like sentences all have the same consequences? Another example might help illustrate the difficulty. Consider now a notion of consequence defined *à la* LP, i.e., as preservation of values that are greater than or equal to **1/2** in $\mathscr{C}$. Here the situation might seem even 'worse': while now not every sentence follows from liar-like sentences, it is still the case that every sentence follows from a revenge or a truth-teller–like sentence, which again might seem counterintuitive. However, these potentially counterintuitive features result from defining consequence using only numerical values, in a semantic framework that also employs equations systems as semantic values. If consequence is defined using also equational values, liar-like, revenge, and truth-teller–like sentences have different sets of consequences.[47] For example, say that φ follows from Γ if:

- either every sentence in $\{\Gamma, \varphi\}$ has a numerical value and whenever every sentence in Γ has value **1**, so does φ,
- or every sentence in $\{\Gamma, \varphi\}$ has an equation system as a value and whenever the system of every sentence in Γ has a solution, so does φ.

To be sure, more refined notions of consequence could (and should) be devised, e.g., to account for premises Γ featuring sentences with both numerical and equational values. However, this simple notion of consequence suffices for present purposes, since it already separates truth-teller–like from revenge cases. More precisely, every sentence now follows from revenge sentences, but not from truth-teller–like sentences. In general, using both numbers *and* equations in defining one's notions on consequence based on $\mathscr{C}$ would make it possible to determine which forms of reasoning are valid for every kind of paradox, distinguishing the consequences of liar-like, truth-teller–like and revenge sentences. I plan to investigate the notions of consequence definable within the semantic framework of the canonical evaluation in future work.[48]

46 For LP see Priest (1979), for ST and TS see Cobreros *et al.,* (2012). See Chemla, Égré, & Spector (2017), Chemla & Égré (2019) for a systematic discussion of many-valued consequence relations.

47 I am indebted to Emmanuel Chemla for suggesting to consider equation systems in the definition of consequence.

48 One might argue that defining a notion of consequence using only one evaluation function is not sufficiently general (even though this seems to be the approach adopted, e.g., in Field (2003, 2008)). However, one can define notions of consequences via sets of evaluations that properly extend the canonical evaluation in its assignments of numerical values. Just like one obtains nonminimal Kripkean fixed points by assigning value **1** or **0** to truth-teller–like sentences, one obtains nonminimal, *quasi-canonical evaluations* that assign numerical values to truth-teller–like sentences. In addition, the algebraic structure determined by quasi-canonical evaluations has several features in common with the structure of Kripkean fixed points (e.g., there are maximal evaluations, intrinsic evaluations, and more). While quasi-canonical evaluations do not provide

– Consider an object-language featuring a predicate for the canonical evaluation itself (such a predicate would be a partial naïve truth predicate, as per Proposition 4.26 and Lemmas 4.27 and 4.28), and now consider a revenge sentence $\rho_{\mathscr{C}}$ equivalent to the claim that $\rho_{\mathscr{C}}$ has a canonical value different from **1**. How is such a $\rho_{\mathscr{C}}$ to be treated by the canonical evaluation (for the language in which $\rho_{\mathscr{C}}$ is formulated)? The answer is quite clear: with an unsolvable equation system. If $\rho_{\mathscr{C}}$ forces evaluation conditions that are impossible to satisfy—having simultaneously a value identical to and different from **1** – then it is a revenge sentence, and it should be evaluated as such. One might object that this is an undesirable result: since $\mathscr{C}(\rho_{\mathscr{C}})$ is an unsolvable equation system, then $\mathscr{C}(\rho_{\mathscr{C}})$ is different from **1**, just as the sentence $\rho_{\mathscr{C}}$ says, so it should actually receive value **1**, which is impossible. However, the objection is far from being devastating: it simply shows that the canonical evaluation is *itself* subject to revenge paradoxes—and this is to be expected, since minimally expressive semantic theories are subject to revenge. The advantage of the canonical evaluation is that it treats *its own revenge sentences* just as it treats every other object-linguistic revenge sentences, namely by assigning them unsolvable equation systems.
– What about axiomatising the semantic theory provided by the canonical evaluation? It is clear that no axiomatisation that is *adequate* in the sense of Fischer, Halbach, Kriener, & Stern (2015) is available, for reasons of computational complexity (the truth-set determined by $\mathscr{C}$ exceed the Δ^1_1-complete subsets of the relevant domain). However, one might wonder whether there are nice ways of characterising the computable fragment of $\mathscr{C}$; I plan to explore this issue in future work.

§5. Concluding remarks. The main objective of the present work has been to propose a unified theory of truth and paradox, that is an (idealised) interpretation of a language with a naïve truth predicate that also provides an interpretation of the paradoxical sentences that arise from the combination of self-applicable semantic notions, logical principles, and syntactic mechanisms. The canonical evaluation provides both a theory of naïve truth and a theory of semantic paradoxes (see Proposition 4.26 and Lemmas 4.27 and 4.28). As far as the theory of paradox goes, I have argued that (in addition to 'nonparadoxical' sentences) three main kinds of paradoxical sentences can be distinguished: liar-like sentences, truth-teller–like sentences, and revenge-sentences (§2.1–§2.4 and §3.1–§3.4). The canonical evaluation captures and expresses the distinction between these fundamental types of paradoxical statements:

– Liar-like sentences: they are compatible with exactly one nonclassical numerical value. The canonical evaluation assigns them their nonclassical numerical value.
– Truth-teller–like sentences: they are compatible with more than one numerical value (classical or nonclassical). The canonical evaluation assigns them equation systems with more than one solution.
– Revenge sentences: they are incompatible with any numerical value (classical or nonclassical). The canonical evaluation assigns them equation systems with no solution.

a nice theory of paradoxes (they conflate truth-teller–like sentences with either nonparadoxical or liar-like sentences, just as it happens in Kripke's theory), collections of quasi-canonical evaluations can be used to give more general notions of consequence.

The above classification is robust, in that it is independent from which compositional semantics is selected for the logical vocabulary. Indeed, a variant of the canonical evaluation presented here can be given for any compositional interpretation for the logical vocabulary.

Much work remains to be done in order to understand and account for the semantic paradoxes. Much work remains to be done even in the framework I have introduced here—some possible developments have been outlined in §4.7. I hope to have provided at least a first step towards a unified account of semantic notions and of the paradoxical phenomena they engender.

§6. Appendix: Proofs of the main results.

PROPOSITION 4.19.

(I) *For every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ *and every ordinal* α:

1. *There is exactly one* e^{α}_{φ}.
2. e^{α}_{φ} *is a* function, *i.e., for every* $\mathrm{v} \in \mathrm{N}_{\varphi}$ *and every* $\mathbf{v_0}, \mathbf{v_1} \in \mathsf{V}_{\text{Ł}} \cup \mathcal{P}(\mathsf{E}_{\text{Ł}})$:
$$\textit{if } \langle \mathrm{v}, \mathbf{v_0}\rangle \in e^{\alpha}_{\varphi} \textit{ and } \langle \mathrm{v}, \mathbf{v_1}\rangle \in e^{\alpha}_{\varphi}, \textit{ then } \mathbf{v_0} = \mathbf{v_1}.$$
I write '$e^{\alpha}_{\varphi}(\mathrm{v}) = \mathbf{v}$*' if* $\langle \mathrm{v}, \mathbf{v}\rangle \in e^{\alpha}_{\varphi}$*, and '*$e^{\alpha}_{\varphi} = e^{\beta}_{\varphi}$*' if, for every* $\mathrm{v} \in \mathrm{N}_{\varphi}$, $e^{\alpha}_{\varphi}(\mathrm{v}) = e^{\beta}_{\varphi}(\mathrm{v})$.
3. *For every* $\mathrm{v} \in \mathrm{N}_{\varphi}$*, if* $e^{\alpha}_{\varphi}(\mathrm{v}) = \mathbf{k}$ *for a* $\mathbf{k} \in \mathsf{V}_{\text{Ł}}$*, then for every* $\beta > \alpha$, $e^{\beta}_{\varphi}(\mathrm{v}) = \mathbf{k}$.
4. *For every* $\mathrm{v} \in \mathrm{N}_{\varphi}$*, if* $e^{\alpha}_{\varphi}(\mathrm{v}) = \mathbf{E}$ *for* $\mathbf{E} \in \mathcal{P}(\mathsf{E}_{\text{Ł}})$*, then for every* $\beta > \alpha$*, if there is no* $\mathbf{k} \in \mathsf{V}_{\text{Ł}}$ *s.t.* $e^{\beta}_{\varphi}(\mathrm{v}) = \mathbf{k}$*, then* $e^{\beta}_{\varphi}(\mathrm{v}) = \mathbf{E}$.

(II) *There is* exactly one *ordinal* δ_0 *s.t. for every* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, $e^{\delta_0}_{\varphi}$ *is a* fixed point *of the functions* e^{α}_{φ}*, i.e., for every* $\delta \geq \delta_0$ *and* $\varphi \in \mathcal{L}_{\mathsf{Tr}}$:
$$e^{\delta_0}_{\varphi} = e^{\delta}_{\varphi}$$
I indicate $e^{\delta_0}_{\varphi}$ *simply as* e_{φ}.

Proof. Ad (I), let φ be any $\mathcal{L}_{\mathsf{Tr}}$-sentence. Item 1 is immediate. Item 2 follows from the next lemma:

LEMMA 6.1. *For every* $\varphi \in \mathcal{L}^{\rightarrow}_{T}$*, every ordinal* α*, and every* $\mathrm{v} \in \mathrm{N}_{\varphi}$:

– *There are* at most two $\mathbf{v_0}, \mathbf{v_1} \in \mathsf{V}_{\text{Ł}} \cup \mathcal{P}(\mathsf{E}_{\text{Ł}})$ *s.t.* $\langle \mathrm{v}, \mathbf{v_0}\rangle, \langle \mathrm{v}, \mathbf{v_1}\rangle \in \Psi_{\varphi}(e^{\alpha}_{\varphi})$.
– *If* $\langle \mathrm{v}, \mathbf{v_0}\rangle, \langle \mathrm{v}, \mathbf{v_1}\rangle \in \Phi_{\varphi}(e^{\alpha}_{\varphi})$*, then one of* $\mathbf{v_0}$ *and* $\mathbf{v_1}$ *is in* $\mathcal{P}(\mathsf{E}_{\text{Ł}})$ *and the other one is in* $\mathsf{V}_{\text{Ł}}$.

Proof. If $\alpha = 0$, the result is trivial. For the successor case, assume the claim up to α, and let $\mathrm{v} \in \mathrm{N}_{\varphi}$ be s.t. $\langle \mathrm{v}, \mathbf{v_0}\rangle, \langle \mathrm{v}, \mathbf{v_1}\rangle \in \Psi_{\varphi}(e^{\alpha}_{\varphi})$, for two distinct $\mathbf{v_0}, \mathbf{v_1} \in \mathsf{V}_{\text{Ł}} \cup \mathcal{P}(\mathsf{E}_{\text{Ł}})$. By IH, for any such $\mathrm{v} \in \mathrm{N}_{\varphi}$, $\mathbf{v_0}$ and $\mathbf{v_1}$ are the only two elements of $\mathsf{V}_{\text{Ł}} \cup \mathcal{P}(\mathsf{E}_{\text{Ł}})$ s.t. $\langle \mathrm{v}, \mathbf{vl_i}\rangle, \langle \mathrm{v}, \mathbf{vl_j}\rangle \in \Psi_{\varphi}(e^{\alpha}_{\varphi})$; moreover, one of them is in $\mathcal{P}(\mathsf{E}_{\text{Ł}})$ (say $\mathbf{v_0}$), and the other is in $\mathsf{V}_{\text{Ł}}$ (say $\mathbf{vl_1}$). By Definition 4.18,
$$\begin{aligned} e^{\alpha+1}_{\varphi} := \Psi_{\varphi}(e^{\alpha}_{\varphi}) \setminus \{\langle x, y\rangle \in \mathrm{N}_{\varphi} \times \mathcal{P}(\mathsf{E}_{\text{Ł}}) \mid \langle x, y\rangle \in \Psi_{\varphi}(e^{\alpha}_{\varphi}) \\ \text{and there is a } \mathbf{k} \in \mathsf{V}_{\text{Ł}} \text{ s.t. } \langle x, \mathbf{k}\rangle \in \Psi_{\varphi}(e^{\alpha}_{\varphi})\} \end{aligned}$$

so $e_\varphi^{\alpha+1}$ is a function, where $\langle \mathtt{v}, \mathbf{v_0}\rangle \in e_\varphi^{\alpha+1}$, while $\langle \mathtt{v}, \mathbf{v_1}\rangle \notin e_\varphi^{\alpha+1}$. Then, in order to show the claim for $\Psi_\varphi(e_\varphi^{\alpha+1})$, we only have to consider all the possible outcomes of an application of Ψ_φ, namely of the clauses of Definition 4.14, to $e_\varphi^{\alpha+1}$. However, since $e_\varphi^{\alpha+1}$ is a function, there are at most two clauses of Definition 4.14 that can apply simultaneously to it, namely:

(a) Clauses 4 and 7, 4 and 9, or 5 and 8: every such pair of clauses yields the same values in $\mathsf{V}_{\text{Ł}}$.[49]

(b) Exactly one of clauses 11–14 and clause 15.[50] If clause 15 and one of clauses 11–14 applies to $e_\varphi^{\alpha+1}$, exactly two pairs obtain, $\langle \mathtt{v}, \mathbf{v_2}\rangle$ and $\langle \mathtt{v}, \mathbf{v_3}\rangle$, where $\mathbf{v_1} \in \mathcal{P}(\mathsf{E}_{\text{Ł}})$ and $\mathbf{v_3} \in \mathsf{V}_{\text{Ł}}$ (since exactly one Łukasiewicz equation system and one numerical value obtain), as desired.

The limit case is straightforward. (Lemma 6.1) □

If there are less than two values assigned to every node by $\Psi_\varphi(e_\varphi^\alpha)$, the claim is immediate. If there are at least two $\mathbf{v_0}, \mathbf{v_1} \in \mathsf{V}_{\text{Ł}} \cup \mathcal{P}(\mathsf{E}_{\text{Ł}})$ assigned to $\mathtt{v}$ by $\Psi_\varphi(e_\varphi^\alpha)$, by Lemma 6.1 they are the only distinct values, one of them is in $\mathcal{P}(\mathsf{E}_{\text{Ł}})$ and the other one is in $\mathsf{V}_{\text{Ł}}$. By Definition 4.18, $\langle \mathtt{v}, \mathbf{v_0}\rangle$ is removed in the revision step (where $\mathbf{v_0} \in \mathcal{P}(\mathsf{E}_{\text{Ł}})$), leaving only $\langle \mathtt{v}, \mathbf{v_1}\rangle$ in $e_\varphi^{\alpha+1}$. The limit case is immediate.

Item 3 of the Proposition follows from the following lemma:

LEMMA 6.2. *For every* $\mathtt{v} \in \mathtt{N}_\varphi$, *if* $e_\varphi^\alpha(\mathtt{v}) = \mathbf{k}$ *for* $\mathbf{k} \in \mathsf{V}_{\text{Ł}}$, *then* $e_\varphi^{\alpha+1}(\mathtt{v}) = \mathbf{k}$.

Proof. If α is 0, the claim is trivial. Let α be $\alpha_0 + 1$. Then, exactly one clause of Definition 4.14 amongst 2–3, 4–10, and 15 applies to e_φ^α.[51] We reason by cases (I only do a few examples.):

– If clause 2 or 3 applies, the claim is immediate, since it applies at every ordinal.
– If clause 8 applies, then $\alpha_0 \geq 1$, $\mathtt{L}_\varphi(\mathtt{v}) = \psi \rightarrow \chi$, and there are $\mathtt{v_m}, \mathtt{v_n}$ in $\mathtt{N}_\varphi$, immediate successors of $\mathtt{v}$, s.t. $\mathtt{L}_\varphi(\mathtt{v_m}) = \psi$ and $\mathtt{L}_\varphi(\mathtt{v_n}) = \chi$. By IH, Definitions 4.14 and 4.18:
$$e_\varphi^{\alpha_0}(\mathtt{v_m}) = \mathbf{i}, \qquad e_\varphi^{\alpha_0}(\mathtt{v_n}) = \mathbf{j},$$
$$e_\varphi^{\alpha_0+1}(\mathtt{v}) = \mathbf{k} = \min[\mathbf{1}, (\mathbf{1} - \mathbf{i} + \mathbf{j})].$$
By IH $e_\varphi^{\alpha_0+1}(\mathtt{v_m}) = \mathbf{i}$ and $e_\varphi^{\alpha_0+1}(\mathtt{v_n}) = \mathbf{j}$, so the conditions of clause 7 are satisfied for $\alpha + 1$, and
$$e_\varphi^{\alpha_0+2}(\mathtt{v}) = e_\varphi^{\alpha+1}(\mathtt{v}) = \mathbf{k} = \min[\mathbf{1}, (\mathbf{1} - \mathbf{i} + \mathbf{j})].$$
– If clause 16 applies, the claim is immediate by the IH and our requirement of the existence of *unique* solutions for equation systems.

The case where α is a limit is similar to the successor case. (Lemma 6.2) □

[49] E.g., if a simple point $\mathtt{v}$ is labelled with $\psi \wedge \chi$, both its immediate successors have a numerical value, and at least one such value is $\mathbf{0}$, then both clauses 4 and 7 apply, assigning value $\mathbf{0}$ to $\mathtt{v}$.

[50] In other words, one of the clauses to assign an equation to a node (11–14) and the clause to solve a given equation system, assigning the corresponding numerical values to nodes (clause 15).

[51] For simplicity, I am ignoring the case of two clauses applying to the same node yielding the same numerical value (as explained in the case (a) of the proof of Lemma 6.1).

Suppose now that $e_{\varphi}^{\alpha}(\mathrm{v}) = \mathbf{k}$ for a $\mathbf{k} \in \mathsf{V}_{\L}$, and let $\beta \geq \alpha$. If $\beta = \alpha$, the claim is trivially true. Suppose that $\beta > \alpha$. If β is $\beta_0 + 1$, assume the claim up to β_0 as IH. Therefore:

$$e_{\varphi}^{\alpha}(\mathrm{v}) = \mathbf{k} = e_{\varphi}^{\beta_0}(\mathrm{v}) = e_{\varphi}^{\beta_0+1}(\mathrm{v}) = e_{\varphi}^{\beta}(\mathrm{v}),$$

by assumption, IH, Lemma 6.2, and the definition of β, respectively. Let β be a limit and assume the claim as IH for every β_0 such that $\alpha \leq \beta_0 < \beta$. For every such every β_0:

$$e_{\varphi}^{\alpha}(\mathrm{v}) = \mathbf{k} = e_{\varphi}^{\beta_0}(\mathrm{v}) = e_{\varphi}^{\beta}(\mathrm{v}),$$

by assumption, IH, and Definition 4.18, respectively.

Item 4 of the Proposition follows from the following lemma:

LEMMA 6.3. *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ and every ordinal α, if there is an $\mathbf{E} \in \mathcal{P}(\mathsf{E}_{\L})$ s.t. $e_{\varphi}^{\alpha}(\mathrm{v}) = \mathbf{E}$ and there is no $\mathbf{k} \in \mathsf{V}_{\L}$ s.t. $e_{\varphi}^{\alpha+1}(\mathrm{v}) = \mathbf{k}$, then $e_{\varphi}^{\alpha+1}(\mathrm{v}) = \mathbf{E}$.*

Proof. If α is $\alpha_0 + 1$, then first exactly one clause of Definition 4.14 amongst 11–14 applies to $e_{\varphi}^{\alpha_0}$, yielding that $e_{\varphi}^{\alpha}(\mathrm{v}) = \mathbf{E}$, for $\mathbf{E} \in \mathcal{P}(\mathsf{E}_{\L})$. Assume the claim as IH up to α. I only do one example. Suppose clause 12 applies and $\mathrm{L}_{\varphi}(\mathrm{v}) = \psi \to \chi$. Then $\alpha_0 \geq 1$ and either v has one immediate successor and it loops back to a predecessor, or v has two immediate successors $\mathrm{v_m}$ and $\mathrm{v_n}$, where $\mathrm{L}_{\varphi}(\mathrm{v_m}) = \psi$ and $\mathrm{L}_{\varphi}(\mathrm{v_n}) = \chi$. Suppose the latter obtains. By IH, and Definitions 4.14 and 4.18:

$$\begin{aligned} &e_{\varphi}^{\alpha_0}(\mathrm{v_m}) = \mathbf{v_0}, \qquad\qquad e_{\varphi}^{\alpha_0}(\mathrm{v_n}) = \mathbf{v_1}, \\ &e_{\varphi}^{\alpha_0+1}(\mathrm{v}) = \{\mathbf{u}_{\varphi \to \psi} =_{\L} \mathsf{min}[\mathbf{1}, (\mathbf{1} - \mathbf{s}_{\varphi} + \mathbf{s}_{\psi})]\} \cup \\ &\qquad\qquad \{x \in \mathsf{E}_{\L} \mid x \text{ is assigned by } \mathsf{I}_{\Phi} \text{ to a node with the same loop top as } \mathrm{v}\}. \end{aligned}$$

where $\mathbf{v_0} \in \mathcal{P}(\mathsf{E}_{\L}) \cup (\mathsf{V}_{\L} \setminus \{\mathbf{0}\})$, $\mathbf{vl_1} \in \mathcal{P}(\mathsf{E}_{\L}) \cup (\mathsf{V}_{\L} \setminus \{\mathbf{1}\})$, at least one of $\mathbf{v_0}$ and $\mathbf{v_1}$ is in $\mathcal{P}(\mathsf{E}_{\L})$, and s_m and s_n are the $\mathcal{L}_{\L}$-terms assigned to $\mathrm{v_m}$ and $\mathrm{v_n}$ as per Definition 4.14. Suppose that there are no $\mathbf{i}, \mathbf{j} \in \mathsf{V}_{\L}$ s.t. $e_{\varphi}^{\alpha_0+1}(\mathrm{v_m}) = \mathbf{i}$ and $e_{\varphi}^{\alpha_0+1}(\mathrm{v_n}) = \mathbf{j}$. By IH then $e_{\varphi}^{\alpha_0+1}(\mathrm{v_m}) = \mathbf{v_0}$, $e_{\varphi}^{\alpha_0+1}(\mathrm{v_n}) = \mathbf{v_1}$, and s_m and s_n are assigned to $\mathrm{v_m}$ and $\mathrm{v_n}$ as above. Hence, the conditions of clause 12 are satisfied for $\alpha + 1$, and

$$e_{\varphi}^{\alpha_0+2}(\mathrm{v}) = e_{\varphi}^{\alpha+1}(\mathrm{v}).$$

The case where α is a limit is similar to the successor case. (Lemma 6.3) □

Suppose that $e_{\varphi}^{\alpha}(\mathrm{v}) = \mathbf{E}$ for $\mathbf{E} \in \mathcal{P}(\mathsf{E}_{\L})$, let $\beta \geq \alpha$, and suppose that there is no $\mathbf{k} \in \mathsf{V}_{\L}$ s.t. $e_{\varphi}^{\beta}(\mathrm{v}) = \mathbf{k}$. If $\beta = \alpha$, the claim is is trivially true. Suppose that $\beta > \alpha$. If β is $\beta_0 + 1$, assume the claim up to β_0 as IH. Since we supposed that there is no $\mathbf{k} \in \mathsf{V}_{\L}$ s.t. $e_{\varphi}^{\beta}(\mathrm{v}) = \mathbf{k}$, then (by claim 3 of the Proposition) for every δ s.t. $\alpha \leq \delta < \beta$ there is no $\mathbf{k} \in \mathsf{V}_{\L}$ s.t. $e_{\varphi}^{\delta}(\mathrm{v}) = \mathbf{k}$ either. Therefore:

$$e_{\varphi}^{\alpha}(\mathrm{v}) = \mathbf{E} = e_{\varphi}^{\beta_0}(\mathrm{v}) = e_{\varphi}^{\beta_0+1}(\mathrm{v}) = e_{\varphi}^{\beta}(\mathrm{v}),$$

by assumption, the IH, Lemma 6.3, and the definition of β, respectively. Let β be a limit and assume as IH the claim for every β_0 s.t. $\alpha \leq \beta_0 < \beta$. As above, since we supposed that there is no $\mathbf{k} \in \mathsf{V}_{\L}$ s.t. $e_{\varphi}^{\beta}(\mathrm{v}) = \mathbf{k}$, then (by item 3 of the Proposition) for every δ s.t. $\alpha \leq \delta < \beta$ there is no $\mathbf{k} \in \mathsf{V}_{\L}$ s.t. $e_{\varphi}^{\delta}(\mathrm{v}) = \mathbf{k}$ either. Then, for every δ s.t. $\alpha \leq \delta < \beta$:

$$e_{\varphi}^{\alpha}(\mathrm{v}) = \mathbf{E} = e_{\varphi}^{\delta}(\mathrm{v}) = e_{\varphi}^{\beta}(\mathrm{v})$$

by assumption, the IH, and Definition 4.18, respectively.

To prove claim (II) of the Proposition, given items 3 and 4 of claim (I), it suffices to notice that there are at most $\beth_2$ distinct functions e^{α}_{φ}, since for every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$, $\mathtt{N}_{\varphi}$ is countable and the cardinality of $\mathsf{V}_{\text{Ł}}$ is either finite or $\beth_1$. Therefore, there is a smallest ordinal ζ_0 at which the revision sequence of Definition 4.18 reaches a fixed point for cardinality reasons. □

PROPOSITION 4.22. *For every* $\varphi, \psi \in \mathcal{L}_{\mathsf{Tr}}$, $\mathtt{v} \in \mathtt{N}_{\varphi}$, $\mathtt{w} \in \mathtt{N}_{\psi}$:

$$\textit{if } \mathtt{L}_{\varphi}(\mathtt{v}) = \mathtt{L}_{\psi}(\mathtt{w}), \textit{ then } \langle \mathtt{N}^{\mathtt{v}}_{\varphi}, \mathtt{S}^{\mathtt{v}}_{\varphi} \rangle \cong_l \langle \mathtt{N}^{\mathtt{w}}_{\psi}, \mathtt{S}^{\mathtt{w}}_{\psi} \rangle.$$

Proof. I begin with two preliminary results.

LEMMA 6.4. *For every* $\varphi, \psi \in \mathcal{L}_{\mathsf{Tr}}$, $\mathtt{v} \in \mathtt{N}_{\varphi}$, *and* $\mathtt{w} \in \mathtt{N}_{\psi}$, *if* $\mathtt{L}_{\varphi}(\mathtt{v}) = \mathtt{L}_{\psi}(\mathtt{w})$ *and no looping clause applies to* $\mathtt{v}$ *and* $\mathtt{w}$, *then* $\mathtt{v}$ *and* $\mathtt{w}$ *have the same number of immediate successors, with identical labels.*

Proof. I do only the case $\mathtt{L}_{\varphi}(\mathtt{v}) = \chi \to \sigma = \mathtt{L}_{\psi}(\mathtt{w})$. Since no looping clause applies to $\mathtt{v}$ and $\mathtt{w}$, by Definitions 4.2 and 4.3 the only clause that applies to $\mathtt{v}$ in the construction of $\langle \mathtt{N}_{\varphi}, \mathtt{S}_{\varphi} \rangle$ *as well as* to $\mathtt{w}$ in the construction of $\langle \mathtt{N}_{\psi}, \mathtt{S}_{\psi} \rangle$ is clause (5.1). Therefore, $\mathtt{v}$ and $\mathtt{w}$ have two immediate successors, $\mathtt{v}_0, \mathtt{v}_1$ and $\mathtt{w}_0, \mathtt{w}_1$ respectively, s.t. $\mathtt{L}_{\varphi}(\mathtt{v}_0) = \chi = \mathtt{L}_{\psi}(\mathtt{w}_0)$ and $\mathtt{L}_{\varphi}(\mathtt{v}_1) = \sigma = \mathtt{L}_{\psi}(\mathtt{w}_1)$. (Lemma 6.4) □

LEMMA 6.5. *For every* $\varphi, \psi \in \mathcal{L}_{\mathsf{Tr}}$, $\mathtt{v} \in \mathtt{N}_{\varphi}$, *and* $\mathtt{w} \in \mathtt{N}_{\psi}$, *if* $\mathtt{L}_{\varphi}(\mathtt{v}) = \mathtt{L}_{\psi}(\mathtt{w})$ *and there is a* $\mathtt{v}_0 \in \mathtt{Pred}_{\varphi}(\mathtt{v})$ *s.t. a looping clause applies to* $\mathtt{v}$ *and* $\mathtt{v}_0$ *but there is no* $\mathtt{w}_0 \in \mathtt{Pred}_{\psi}(\mathtt{w})$ *s.t.* $\mathtt{L}_{\varphi}(\mathtt{v}_0) = \mathtt{L}_{\psi}(\mathtt{w}_0)$ *and a looping clause applies to* $\mathtt{w}$ *and* $\mathtt{w}_0$, *then* $\mathtt{w}$ *has a successor* $\mathtt{w}_1$ *s.t.* $\mathtt{L}_{\varphi}(\mathtt{v}_0) = \mathtt{L}_{\psi}(\mathtt{w}_1)$.

Proof. I do only the case $\mathtt{L}_{\varphi}(\mathtt{v}) = \forall x \chi(x) = \mathtt{L}_{\psi}(\mathtt{w})$. If there is a $\mathtt{v}_0 \in \mathtt{Pred}_{\varphi}(\mathtt{v})$ s.t. a looping clause applies to $\mathtt{v}$ and $\mathtt{v}_0$, then by Definition 4.2 such clause is $(7.L)$ and $\mathtt{L}_{\varphi}(\mathtt{v}_0) = \chi(t_n)$ for some term t_n. If there is no $\mathtt{w}_0 \in \mathtt{Pred}_{\psi}(\mathtt{w})$ s.t. $\mathtt{L}_{\varphi}(\mathtt{v}_0) = \mathtt{L}_{\psi}(\mathtt{w}_0)$ and a looping clause applies to $\mathtt{w}$ and $\mathtt{w}_0$, then by Definition 4.2 the only possible clause that applies to $\mathtt{w}$ is (7.1), which yields that $\mathtt{w}$ has an immediate successor $\mathtt{w}_1$ s.t. $\mathtt{L}_{\psi}(\mathtt{w}_1) = \chi(t_n)$, proving the claim. (Corollary 6.5) □

I now turn to the proof of the Proposition. There are three main cases:

Case 1. Let $\mathtt{v}$ be a dead end. By Definition 4.3, $\mathtt{v}$ is either labelled with an atomic $\mathcal{L}$-sentence, or with a sentence $\mathsf{Tr}(t)$ where t does not denote the code of a $\mathcal{L}_{\mathsf{Tr}}$-sentence. Since $\mathtt{L}_{\varphi}(\mathtt{v}) = \mathtt{L}_{\psi}(\mathtt{w})$, $\mathtt{w}$ is also a dead end.

Case 2. Suppose for a contradiction that $\mathtt{v}$ is a simple point but $\mathtt{w}$ is not. As per Case 1, $\mathtt{w}$ is not a dead end, so it belongs to at least one loop $\mathtt{L}$, of the following form:[52]

$$\mathtt{L} = \Big\langle \langle \mathtt{w}_1, \mathtt{w}_2 \rangle, \ldots, \langle \mathtt{w}_\mathtt{i}, \mathtt{w} \rangle, \langle \mathtt{w}, \mathtt{w}_{\mathtt{i}+1} \rangle, \ldots, \langle \mathtt{w}_\mathtt{j}, \mathtt{w}_{\mathtt{j}+1} \rangle, \langle \mathtt{w}_{\mathtt{j}+1}, \mathtt{w}_1 \rangle \Big\rangle$$

where a looping clause applies to $\mathtt{w}_{\mathtt{j}+1}$ (and possibly other nodes). Since $\mathtt{L}_{\varphi}(\mathtt{v}) = \mathtt{L}_{\psi}(\mathtt{w})$, by Lemmas 6.4 and 6.5, $\mathtt{v}$ and $\mathtt{w}$ can have many immediate successors, possibly infinitely many, but in particular $\mathtt{v}$ has an immediate successor $\mathtt{v}_1$ s.t. $\mathtt{L}_{\psi}(\mathtt{w}_{\mathtt{i}+1}) = \mathtt{L}_{\varphi}(\mathtt{v}_1)$. But since we assumed that $\mathtt{v}$ is a simple point, $\mathtt{v}_1$ cannot loop back to $\mathtt{v}$ nor to any of its predecessors. Moreover, it is not the case that $\mathtt{v}_1$ *only* loops back to itself, because otherwise $\mathtt{w}_{\mathtt{i}+1}$ would do that as well (by the above Lemma and Corollary). So, $\mathtt{v}_1$ is also a

[52] In this proof, I suppose for the sake of readability and without loss of generality that the nodes in the loops and paths to be mentioned are enumerated progressively.

simple point. Now we apply to $\mathtt{v}_1$ and $\mathtt{w}_{i+1}$ the same reasoning that we applied to $\mathtt{v}$ and $\mathtt{w}$: using again Lemmas 6.4 and 6.5 on $\mathtt{v}_1$ and $\mathtt{w}_{i+1}$ we conclude that, amongst possibly many others, $\mathtt{v}_1$ has at least one successor labelled as the immediate successor of $\mathtt{w}_{i+1}$ that belongs to $\mathtt{L}$, and that such successor of $\mathtt{v}_1$ is a simple point. Proceeding in this way, by repeated applications of Lemmas 6.4 and 6.5, we generate a path $\mathtt{P} \in \mathtt{R}_\varphi(\mathtt{v})$ s.t., as set of ordered pairs of sentences (labels), matches exactly the loop $\mathtt{L}$.[53] So, $\mathtt{P}$ has the form:

$$\mathtt{P} = \Big\langle \langle \mathtt{v}, \mathtt{v}_1 \rangle, \langle \mathtt{v}_1, \mathtt{v}_2 \rangle, \ldots, \langle \mathtt{v}_k, \mathtt{v}_{k+1} \rangle \Big\rangle$$

where $\mathtt{L}_\varphi(\mathtt{v}) = \mathtt{L}_\psi(\mathtt{w}), \ldots, \mathtt{L}_\varphi(\mathtt{v}_{k+1}) = \mathtt{L}_\psi(\mathtt{w}_i)$ (the dots stand for the nodes in $\mathtt{L}$) and $j = k$. However, by our assumption that $\mathtt{v}$ is a simple point, a looping clause never applies to a node in $\mathtt{P}$ and $\mathtt{v}$ or one of its predecessors. But, by Lemma 6.5 and Definition 4.3, a looping clause applies at least once to $\mathtt{v}_{k+1}$ that loops back to $\mathtt{v}$. Contradiction. So, if $\mathtt{v}$ is a simple point, also $\mathtt{w}$ is a simple point. They have the same label by assumption and, as shown, they have the same number of immediate successors, labelled with the same sentences, by Lemma 6.4.

Case 3. Let $\mathtt{v}$ belong to at least one loop $\mathtt{P}^*$ and suppose for a contradiction that there is no loop $\mathtt{P} \in \mathtt{R}_\psi(\mathtt{w})$ s.t. $\mathtt{P}^*$ and $\mathtt{P}$ have the same number of nodes and identical labels for edges (within $\mathtt{L}_\varphi$ and $\mathtt{L}_\psi$, respectively).[54] $\mathtt{P}^*$ is a finite set of ordered pairs of labelled nodes of the following form:

$$\mathtt{P}^* = \Big\langle \langle \mathtt{v}_1, \mathtt{v}_2 \rangle, \ldots, \langle \mathtt{v}_i, \mathtt{v} \rangle, \langle \mathtt{v}, \mathtt{v}_{i+1} \rangle, \ldots, \langle \mathtt{v}_k, \mathtt{v}_{k+1} \rangle, \langle \mathtt{v}_{k+1}, \mathtt{v}_1 \rangle \Big\rangle$$

where a looping clause applies to $\mathtt{v}_{k+1}$ that loops back to $\mathtt{v}_1$. By our supposition, there is no loop in $\mathtt{R}_\psi(\mathtt{w})$ that is loop-isomorphic to $\mathtt{P}^*$. By cases 1 and 2, $\mathtt{w}$ is in at least one loop. If $\mathtt{w}$ is not contained in any loop which is loop-isomorphic to $\mathtt{P}^*$, then a looping clause applies to $\mathtt{w}$ or to one of its successors (within $\mathtt{R}_\psi(\mathtt{w})$) that are labelled as a node in $\mathtt{P}^*$, but not to such node in $\mathtt{P}^*$, or *vice versa*. Let $\mathtt{w}_{p+1}$ be the first node in $\mathtt{R}_\psi(\mathtt{w})$ s.t. the first case is given (otherwise, it is dual). So, there is a path $\mathtt{P}$

$$\Big\langle \langle \mathtt{w}, \mathtt{w}_{n+1} \rangle, \ldots, \langle \mathtt{w}_p, \mathtt{w}_{p+1} \rangle, \langle \mathtt{w}_{p+1}, \mathtt{w}_m \rangle \Big\rangle = \mathtt{P} \in \mathtt{R}_\psi(\mathtt{w})$$

s.t. $\mathtt{L}_\psi(\mathtt{w}_{n+1}) = \mathtt{L}_\varphi(\mathtt{v}_{i+1}), \ldots, \mathtt{L}_\psi(\mathtt{w}_p) = \mathtt{L}_\varphi(\mathtt{v}_{i+(p-n)}), \mathtt{L}_\psi(\mathtt{w}_{p+1}) = \mathtt{L}_\varphi(\mathtt{v}_{i+(p-n)+1})$, by Lemmas 6.4 and 6.5. A looping clause applies to $\mathtt{w}_{p+1}$ and some $\mathtt{w}_m$ labelled as $\mathtt{v}_{i+(p-n)+2}$ (by our assumption), but it does not apply to $\mathtt{v}_{i+(p-n)+1}$ and $\mathtt{v}_{i+(p-n)+2}$ (by our assumption). Then $\mathtt{v}_{i+(p-n)+1}$ is between $\mathtt{v}_{i+1}$ and $\mathtt{v}_k$ in $\mathtt{P}^*$, so there is an index l s.t. $\mathtt{v}_{i+(p-n)+1}$ is $\mathtt{v}_l$. Therefore, $\mathtt{P}^*$ looks as follows:

$$\mathtt{P}^* = \Big\langle \langle \mathtt{v}_1, \mathtt{v}_2 \rangle, \ldots, \langle \mathtt{v}_i, \mathtt{v} \rangle, \langle \mathtt{v}, \mathtt{v}_{i+1} \rangle, \ldots, \langle \mathtt{v}_l, \mathtt{v}_{l+1} \rangle, \ldots, \langle \mathtt{v}_k, \mathtt{v}_{k+1} \rangle, \langle \mathtt{v}_{k+1}, \mathtt{v}_1 \rangle \Big\rangle.$$

Since a looping clause applies to $\mathtt{w}_{p+1}$, there is a path $\mathtt{P}_1$ s.t.

$$\mathtt{P} \subsetneq \Big\langle \langle \mathtt{w}_m, \mathtt{w}_{m+1} \rangle, \ldots, \langle \mathtt{w}, \mathtt{w}_{n+1} \rangle, \ldots, \langle \mathtt{w}_p, \mathtt{w}_{p+1} \rangle, \langle \mathtt{w}_{p+1}, \mathtt{w}_m \rangle \Big\rangle = \mathtt{P}_1 \in \mathtt{R}_\psi(\mathtt{w})$$

53 Corollary 6.5 is used in generating the successor of $\mathtt{v}$ (within $\mathtt{P}$) labelled as $\mathtt{w}_{j+1}$.

54 An anonymous referee has described the proof strategy employed in this case of the demonstration as the application of a kind of pigeonhole principle. In fact, I show that, trying to systematically falsify the claim of (this case of) the Proposition and constructing all the possible paths in $\mathtt{R}_\psi(\mathtt{w})$ that start from $\mathtt{w}$, one can derive the existence of a path that is loop-isomorphic to $\mathtt{P}^*$.

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

where $\mathtt{L}_\psi(\mathtt{w_m}) = \mathtt{L}_\varphi(\mathtt{v_{l+1}})$ and $\mathtt{L}_\psi(\mathtt{w_{m+1}}) = \mathtt{L}_\varphi(\mathtt{v_{l+2}})$. Our assumption, however, dictates that there is no path within $\mathtt{R}_\psi(\mathtt{w})$ that is loop-isomorphic to $\mathtt{P}^*$. So, any path within $\mathtt{R}_\psi(\mathtt{w})$ that contains nodes labelled as $\mathtt{L}_\psi(\langle \mathtt{w_m}, \mathtt{w_{m+1}}\rangle \cup \mathtt{P})$, with the same order of labels as $\mathtt{P}^*$, is not loop-isomorphic to $\mathtt{P}^*$.[55] There are at most countably many such paths $\mathtt{P_{1.1}}, \mathtt{P_{1.2}}, \ldots$ and $\mathtt{P_1}$ is one of them. Considering all the (at most countably infinitely many) possible cases, we have:

$$\Big\langle \langle \mathtt{w_m}, \mathtt{w_{m+1}}\rangle, \sim_1, \langle \mathtt{w}, \mathtt{w_{n+1}}\rangle, \ldots, \langle \mathtt{w_{p+1}}, \mathtt{w_m}\rangle \Big\rangle = \mathtt{P_{1.1}} \in \mathtt{R}_\psi(\mathtt{w})$$

$$\vdots$$

$$\Big\langle \langle \mathtt{w_m}, \mathtt{w_{m+1}}\rangle, \sim_j, \langle \mathtt{w^j}, \mathtt{w^j_{n+1}}\rangle, \ldots, \langle \mathtt{w^j_{p+1}}, \mathtt{w_m}\rangle \Big\rangle = \mathtt{P_{1.j}} \in \mathtt{R}_\psi(\mathtt{w})$$

$$\vdots$$

where $\sim_j$ indicates some path *different* in labels from $\mathtt{L}_\varphi(\mathtt{P}^*) \setminus [\mathtt{L}_\psi(\langle \mathtt{w_m}, \mathtt{w_{m+1}}\rangle) \cup \bigcap_i \mathtt{L}_\psi(\mathtt{P_{1.i}})]$ and in their order. Moreover, for $i \in \omega$, $\mathtt{L}_\psi(\mathtt{w}) = \mathtt{L}_\psi(\mathtt{w^i})$, $\mathtt{L}_\psi(\mathtt{w_{n+1}}) = \mathtt{L}_\psi(\mathtt{w^i_{n+1}})$, …, $\mathtt{L}_\psi(\mathtt{w_{p+1}}) = \mathtt{L}_\psi(\mathtt{w^i_{p+1}})$.

Lemmas 6.4 and 6.5 imply that there is a node in $\mathtt{R}_\psi(\mathtt{w_{m+1}})$ labelled as some member of $\mathtt{P}^*$ s.t. a looping clause applies to that node but not to the corresponding member of $\mathtt{P}^*$, or *vice versa* (otherwise there would be a path loop-isomorphic to $\mathtt{P}^*$ within $\mathtt{R}_\psi(\mathtt{w})$). In terms of the above list, this means that some node in some of the $\mathtt{P_{1.n}}$ in the disagreeing part of the path $\sim_n$ loops back to a predecessor of $\mathtt{w_m}$ (so, such node is a predecessor of all nodes that are successors of $\mathtt{w_m}$).[56]

Take any such path, call it $\mathtt{P_2}$. We have that:

$$\Big\langle \langle \mathtt{w_o}, \mathtt{w_{o+1}}\rangle, \ldots, \langle \mathtt{w_q}, \mathtt{w_{q+1}}\rangle, \langle \mathtt{w_{q+1}}, \mathtt{w_o}\rangle \Big\rangle = \mathtt{P_2} \in \mathtt{R}_\psi(\mathtt{w})$$

for $\mathtt{L}_\psi(\mathtt{w_o}) = \mathtt{L}_\varphi(\text{some node in } \mathtt{P}^*)$, $\mathtt{L}_\psi(\mathtt{w_{o+1}}) = \mathtt{L}_\varphi(\text{its successor labelled as the corresponding node in } \mathtt{P}^*)$ and a looping clause applies to $\mathtt{w_{q+1}}$ and $\mathtt{w_o}$. Clearly $\mathtt{L}_\psi(\mathtt{w_{q+1}}) \neq \mathtt{L}_\psi(\mathtt{w^n_{p+1}})$ for all $n \in \omega$ (by construction and because otherwise no path such as $\mathtt{P_{1.n}}$ could exist). By construction, $\mathtt{w_o}, \mathtt{w_{o+1}}$ are predecessors of any node in $\bigcup_i \mathtt{P_{1.i}}$. Now, we reason as we did after finding the path $\mathtt{P_1}$, deriving the existence of (possibly countably infinitely many) paths $\mathtt{P_{2.1}}, \mathtt{P_{2.2}}, \ldots$ such that:

$$\Big\langle \langle \mathtt{w_o}, \mathtt{w_{o+1}}\rangle, \sim_1, \langle \mathtt{w_q}, \mathtt{w_{q+1}}\rangle, \ldots, \langle \mathtt{w_{q+1}}, \mathtt{w_o}\rangle \Big\rangle = \mathtt{P_{2.1}} \in \mathtt{R}_\psi(\mathtt{w})$$

$$\vdots$$

$$\Big\langle \langle \mathtt{w_o}, \mathtt{w_{o+1}}\rangle, \sim_j, \langle \mathtt{w^j_q}, \mathtt{w^j_{q+1}}\rangle, \ldots, \langle \mathtt{w^j_{q+1}}, \mathtt{w^j_o}\rangle \Big\rangle = \mathtt{P_{2.j}} \in \mathtt{R}_\psi(\mathtt{w})$$

$$\vdots$$

where $\sim_j$ indicates a path that is different in labels and their order from $\mathtt{L}_\varphi(\mathtt{P}^*) \setminus \mathtt{L}_\psi[(\langle \mathtt{w_o}, \mathtt{w_{o+1}}\rangle) \cup \bigcap_i \mathtt{L}_\psi(\mathtt{P_{2.i}})]$.

[55] I write $\mathtt{L}_\psi(\mathtt{P})$ to denote the labels assigned by $\mathtt{L}_\psi$ to the nodes in $\mathtt{P}$.

[56] Such a node exists by Lemmas 6.4 and 6.5 and our assumption.

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

Then we take an arbitrary path $\mathtt{P}_3$, as we did for $\mathtt{P}_2$, i.e., s.t. a looping clause applies to two nodes in it labelled as in $\mathtt{P}^*$, but s.t. (by assumption) the resulting path is not loop-isomorphic to $\mathtt{P}^*$. By construction, the node labelled as a node in $\mathtt{P}^*$ that loops back to a node labelled as a node in $\mathtt{P}^*$ is s.t. no node with the same label loops back to a node labelled as a node in $\mathtt{P}^*$ in $\mathtt{P}_1$ or $\mathtt{P}_2$ (just as above we had that $\mathtt{L}_\psi(\mathtt{w}_{\mathtt{q}+1}) \neq \mathtt{L}_\psi(\mathtt{w}^{\mathtt{n}}_{\mathtt{p}+1})$ for all $n \in \omega$). Moreover, the node in $\mathtt{P}_3$ to which such node loops back is a predecessor of every node in $(\bigcup_i \mathtt{P}_{1.\mathtt{i}} \cup \bigcup_i \mathtt{P}_{2.\mathtt{i}})$ (just as we had that $\mathtt{w}_\mathtt{o}$ and $\mathtt{w}_{\mathtt{o}+1}$ are predecessors of every node in $\bigcup_i \mathtt{P}_{1.\mathtt{i}}$ above).

Proceeding in this way, we derive the existence of more and more paths $\mathtt{P}_4, \mathtt{P}_5, \ldots$ s.t. every $\mathtt{P}_{\mathtt{t}+1}$ in $\mathtt{P}_1, \mathtt{P}_2, \mathtt{P}_3, \mathtt{P}_4, \mathtt{P}_5, \ldots$ is loop-isomorphic to a proper subpath of $\mathtt{P}^*$, but is not loop-isomorphic to $\mathtt{P}^*$ itself. The existence of such paths is guaranteed by iterated applications of Lemmas 6.4 and 6.5, while the ‘asymmetric’ looping back we have seen in the construction of $\mathtt{P}_1, \mathtt{P}_2, \mathtt{P}_3, \ldots$ is forced by our assumption that no such path can ever be loop-isomorphic to $\mathtt{P}^*$: if no node in the path in question looped back asymmetrically, our assumption would be immediately falsified.

The key facts are: (1) for every $\mathtt{P}_{\mathtt{t}+1}$ amongst $\mathtt{P}_1, \mathtt{P}_2, \mathtt{P}_3, \ldots$, a looping clause applies to a node $\mathtt{w}'$ in $\mathtt{P}_{\mathtt{t}+1}$ labelled as a node in $\mathtt{P}^*$ but no node $\mathtt{w}''$ with the same label as $\mathtt{w}'$ in any of $\mathtt{P}_1, \mathtt{P}_2, \mathtt{P}_3, \mathtt{P}_4, \ldots, \mathtt{P}_\mathtt{t}$ is s.t. a looping clause applies to $\mathtt{w}''$ and to a node labelled as a node in $\mathtt{P}^*$ (by construction and to avoid contradiction, as shown above); (2) for every $\mathtt{P}_{\mathtt{t}+1}$ amongst $\mathtt{P}_1, \mathtt{P}_2, \mathtt{P}_3, \ldots$, a node labelled as an element of $\mathtt{P}^*$ loops back to a node which is also labelled as an element of $\mathtt{P}^*$: the latter is a predecessor of every node in $\bigcup_{m \le t}(\bigcup_i \mathtt{P}_{\mathtt{m.i}})$ (by the definition of predecessor and the above construction, which in turn is forced by Lemmas 6.4 and 6.5 together with our assumption).

But $\mathtt{P}^*$ has only *finitely many nodes*, so the process described in key facts (1) and (2) cannot go on forever. Deriving the existence of finitely many $\mathtt{P}_1, \mathtt{P}_2, \mathtt{P}_3, \mathtt{P}_4, \ldots$, we should find a path $\mathtt{P}_{\mathtt{n}+1}$

$$\Big\langle \langle \mathtt{w}_\mathtt{r}, \mathtt{w}_{\mathtt{r}+1}\rangle, \ldots, \langle \mathtt{w}_\mathtt{s}, \mathtt{w}_{\mathtt{s}+1}\rangle, \langle \mathtt{w}_{\mathtt{s}+1}, \mathtt{w}_\mathtt{r}\rangle \Big\rangle = \mathtt{P}_{\mathtt{n}+1}$$

s.t. $\mathtt{L}_\psi(\mathtt{w}_{\mathtt{s}+1})$ is identical to the label of some node between $\mathtt{v}_{\mathtt{i}+2}$ and $\mathtt{v}_{\mathtt{l}-1}$ (by Lemmas 6.4 and 6.5) and a looping clause applies to $\mathtt{w}_{\mathtt{s}+1}$ and $\mathtt{w}_\mathtt{r}$ (by our supposition and key fact (1)). But this cannot be! For, if there were a node labelled as $\mathtt{w}_\mathtt{r}$ with which $\mathtt{w}_{\mathtt{s}+1}$ is in a loop, then there would have already been some node between $\mathtt{w}_{\mathtt{n}+1}$ and $\mathtt{w}_\mathtt{p}$ (included) within $\mathtt{P}_1$ in a loop with it. This is because $\mathtt{w}_\mathtt{r}$ is a predecessor of every node between $\mathtt{w}_{\mathtt{n}+1}$ and $\mathtt{w}_\mathtt{p}$ (included), by key fact (2). So, no path such as $\mathtt{P}_1$ could have existed, because the nodes between $\mathtt{w}_{\mathtt{n}+1}$ and $\mathtt{w}_\mathtt{p}$ are labelled as the nodes between $\mathtt{v}_{\mathtt{i}+2}$ and $\mathtt{v}_{\mathtt{l}-1}$ (included) and $\mathtt{w}_{\mathtt{p}+1}$ would not be the first successor of $\mathtt{w}$ in $\mathtt{R}_\psi(\mathtt{w})$ labelled as a member of a pair in $\mathtt{P}^*$ s.t. the rule ($\mathtt{Loop}$) applies to it but not to the corresponding member of $\mathtt{P}^*$. Then, we derive the existence of the following path

$$\Big\langle \langle \mathtt{w}_\mathtt{s}, \mathtt{w}_{\mathtt{s}+1}\rangle, \langle \mathtt{w}_{\mathtt{s}+1}, \mathtt{w}_{\mathtt{s}+2}\rangle \Big\rangle = \widehat{\mathtt{P}}_1 \in \mathtt{R}_\psi(\mathtt{w})$$

where $\mathtt{L}_\psi(\mathtt{w}_{\mathtt{s}+1}) = \mathtt{L}_\varphi(\mathtt{v}_{\mathtt{i}+2})$ (by Lemmas 6.4 and 6.5) and no looping clause applies to $\mathtt{w}_{\mathtt{s}+1}$ and any other node labelled as in $\mathtt{P}^*$.

We reiterate this reasoning $k+1$ times, observing that a looping clause never applies to $\mathtt{w}_{\mathtt{s}+\mathtt{m}}$ (for $m \le k+1$) and to some node in $\mathtt{R}_\psi(\mathtt{w})$ labelled as a node in $\mathtt{P}^*$, otherwise we would have a contradiction with the existence of the paths $\mathtt{P}_1, \mathtt{P}_2, \mathtt{P}_3, \mathtt{P}_4, \ldots$ derived so far. We therefore obtain the existence of the following paths:

$$\Big\langle \langle \mathtt{w_s}, \mathtt{w_{s+1}} \rangle, \langle \mathtt{w_{s+1}}, \mathtt{w_{s+2}} \rangle, \langle \mathtt{w_{s+2}}, \mathtt{w_{s+3}} \rangle \Big\rangle = \widehat{\mathtt{P}}_2 \in \mathtt{R}_\psi(\mathtt{w})$$

$$\vdots$$

$$\Big\langle \langle \mathtt{w_s}, \mathtt{w_{s+1}} \rangle, \langle \mathtt{w_{s+1}}, \mathtt{w_{s+2}} \rangle, \ldots, \langle \mathtt{w_{s+k}}, \mathtt{w_{s+k+1}} \rangle, \langle \mathtt{w_{s+k+1}}, \mathtt{w_s} \rangle \Big\rangle = \widehat{\mathtt{P}}_\mathtt{k} \in \mathtt{R}_\psi(\mathtt{w}).$$

By construction and Lemmas 6.4 and 6.5, the labels of the edges of $\widehat{\mathtt{P}}_\mathtt{k}$ are pairwise identical to those of $\mathtt{P}^*$, i.e., $\widehat{\mathtt{P}}_\mathtt{k} \cong_l \mathtt{P}^*$. Contradiction. □

PROPOSITION 4.23. *For every* $\varphi, \psi \in \mathcal{L}_{\mathsf{Tr}}$, $\mathtt{v} \in \mathtt{N}_\varphi$, $\mathtt{w} \in \mathtt{N}_\psi$, *and* $\mathbf{v} \in \mathsf{V}_{\text{Ł}} \cup \mathcal{P}(\mathsf{E}_{\text{Ł}})$, *if* $\mathtt{L}_\varphi(\mathtt{v}) = \mathtt{L}_\psi(\mathtt{w})$:

1. *there is an* α *s.t.* $e^\alpha_\varphi(\mathtt{v}) = \mathbf{v}$ *if and only if there is a* β *s.t.* $e^\beta_\psi(\mathtt{w}) = \mathbf{v}$, *and*
2. $e_\varphi(\mathtt{v}) = e_\psi(\mathtt{w})$.

First, I prove a useful lemma.

LEMMA 6.6. *Let* $\varphi, \psi \in \mathcal{L}_{\mathsf{Tr}}$, $\mathtt{v} \in \mathtt{N}_\varphi$, *and* $\mathtt{w} \in \mathtt{N}_\psi$ *be s.t.* $\mathtt{L}_\varphi(\mathtt{v}) = \mathtt{L}_\psi(\mathtt{w})$. *For every dead end or simple point* $\mathtt{v_i} \in \mathtt{N}^\mathtt{v}_\varphi$ *there is a dead end or simple point* $\mathtt{w_j} \in \mathtt{N}^\mathtt{w}_\psi$ *s.t.*:

– $\mathtt{L}_\varphi(\mathtt{v_i}) = \mathtt{L}_\psi(\mathtt{w_j})$,
– $\mathtt{L}_\varphi$(*the immediate predecessor of* $\mathtt{v_i}$) $= \mathtt{L}_\psi$(*the immediate predecessor of* $\mathtt{w_j}$).

Proof. For $\mathtt{v_i}$ as in the lemma, the immediate predecessor of $\mathtt{v_i}$, call it $\mathtt{v_i}'$, is in $\mathtt{N}^\mathtt{v}_\varphi$. Since by Proposition 4.22 $\langle \mathtt{N}^\mathtt{v}_\varphi, \mathtt{S}^\mathtt{v}_\varphi \rangle \cong_l \langle \mathtt{N}^\mathtt{w}_\psi, \mathtt{S}^\mathtt{w}_\psi \rangle$, there is a $\mathtt{w}' \in \mathtt{N}^\mathtt{w}_\psi$ s.t. $\mathtt{L}_\varphi(\mathtt{v_i}') = \mathtt{L}_\psi(\mathtt{w}')$. The existence of a node $\mathtt{w_j}$ as in the statement of the lemma follows immediately from case 2 of the proof of Proposition 4.22—in fact, if $\mathtt{w_j}$ is not in a loop, its being a simple point or a dead end depends only on its label. (Lemma 6.6) □

Proof sketch of the Proposition. Let φ, ψ, $\mathtt{v}$, and $\mathtt{w}$ be as in the statement of the proposition. I only do the left-to-right direction. Let there be an α s.t. $e^\alpha_\varphi(\mathtt{v}) = \mathbf{v}$. By Definition 4.14, the only nodes of $\mathtt{N}_\varphi$ used in constructing $e^\alpha_\varphi(\mathtt{v})$ are those in $\mathtt{N}^\mathtt{v}_\varphi$. Since $\mathtt{L}_\varphi(\mathtt{v}) = \mathtt{L}_\psi(\mathtt{w})$, by Proposition 4.22 $\langle \mathtt{N}^\mathtt{v}_\varphi, \mathtt{S}^\mathtt{v}_\varphi \rangle \cong_l \langle \mathtt{N}^\mathtt{w}_\psi, \mathtt{S}^\mathtt{w}_\psi \rangle$. As a consequence:

(i) Dead ends of $\langle \mathtt{N}^\mathtt{v}_\varphi, \mathtt{S}^\mathtt{v}_\varphi \rangle$ are mapped to dead ends of $\langle \mathtt{N}^\mathtt{w}_\psi, \mathtt{S}^\mathtt{w}_\psi \rangle$ with identical labels, and *vice versa*.
(ii) Simple points of $\langle \mathtt{N}^\mathtt{v}_\varphi, \mathtt{S}^\mathtt{v}_\varphi \rangle$ are mapped to simple points of $\langle \mathtt{N}^\mathtt{w}_\psi, \mathtt{S}^\mathtt{w}_\psi \rangle$ with identical labels, and by Lemma 6.6 every path made of simple points within $\langle \mathtt{N}^\mathtt{v}_\varphi, \mathtt{S}^\mathtt{v}_\varphi \rangle$ (with the only possible exception of starting with a looping point or ending in a dead end) is reconstructed, identical in order and labels, within $\langle \mathtt{N}^\mathtt{w}_\psi, \mathtt{S}^\mathtt{w}_\psi \rangle$, and *vice versa*.
(iii) Every loop within $\langle \mathtt{N}^\mathtt{v}_\varphi, \mathtt{S}^\mathtt{v}_\varphi \rangle$ is mapped to a loop of $\langle \mathtt{N}^\mathtt{w}_\psi, \mathtt{S}^\mathtt{w}_\psi \rangle$, with identical number of nodes and labels, and *vice versa*.

It is easy to show that nodes with identical labels are assigned the same equations un the first inductive construction, encoded by the minimal fixed point I_Φ. It follows that:

– If $\mathtt{v}$ is a dead end, $e^1_\varphi(\mathtt{v}) = \mathbf{k} = e^1_\psi(\mathtt{w})$, by item (i) (where $\mathbf{k} = \mathbf{0}$ or $\mathbf{k} = \mathbf{1}$).
– If $\mathtt{v}$ is a looping leaf, then $\mathtt{v}$ and $\mathtt{w}$ belong to loop-isomorphic loops and so are assigned the same equation system, by item (iii). And identical equation systems have identical solutions or no solutions.

- If $\mathtt{v}$ is a simple point and $e^{\alpha}_{\varphi}(\mathtt{v})$ has a value, then one or more of the successors of $\mathtt{v}$ has already a value in $\mathsf{V}_{\L} \cup \mathcal{P}(\mathsf{E}_{\L})$ at an ordinal $\gamma < \alpha$;[57] by item (ii), then for some ordinal β the function $e^{\beta}_{\psi}(\mathtt{w})$ has the same value, for the very same reason.
- Finally, the evaluation functions move towards the root node either via paths of simple points or via loops that are 'closer' to the root than those already evaluated. In both cases (as seen with the two previous points), e_{φ} and e_{ψ} give identical values to nodes with identical labels. □

PROPOSITION 4.25. *For every $\varphi, \psi \in \mathcal{L}_{\mathsf{Tr}}$ and $\chi(x) \in \mathsf{For}_{\mathcal{L}_{\mathsf{Tr}}}$, the following holds ($\Longleftrightarrow$ stands for the meta-linguistic 'if and only if', and $\Longrightarrow$ for 'if ... then') :*

$$\begin{aligned}
\mathscr{C}(\neg\varphi) = \mathbf{1} &\iff \mathscr{C}(\varphi) = \mathbf{0}\\
\mathscr{C}(\varphi \wedge \psi) = \mathbf{1} &\iff \mathscr{C}(\varphi) = \mathbf{1} \textit{ and } \mathscr{C}(\psi) = \mathbf{1}\\
\mathscr{C}(\varphi \to \psi) = \mathbf{1} &\iff \mathscr{C}(\varphi) = \mathbf{0},\\
&\qquad \textit{or } \mathscr{C}(\psi) = \mathbf{1},\\
&\qquad \textit{or } \mathscr{C}(\varphi) = \mathbf{j}, \mathscr{C}(\psi) = \mathbf{k}, \textit{ and } \mathbf{j} \le \mathbf{k}\\
\mathscr{C}(\forall x \chi(x)) = \mathbf{1} &\iff \mathscr{C}(\chi(t_k)) = \mathbf{1} \textit{ for all } t_k \in \mathsf{CTer}_{\mathcal{L}_{\mathsf{Tr}}}\\
\mathscr{C}(\mathsf{Tr}(\ulcorner\varphi\urcorner)) = \mathbf{1} &\iff \mathscr{C}(\varphi) = \mathbf{1}.
\end{aligned}$$

In addition, modus ponens *holds for the canonical evaluation*:

$$\mathscr{C}(\varphi) = \mathbf{1} \textit{ and } \mathscr{C}(\varphi \to \psi) = \mathbf{1} \implies \mathscr{C}(\psi) = \mathbf{1}.$$

Proof sketch. I do only one case. Let $\mathscr{C}(\varphi \to \psi) = \mathbf{1}$, $\mathscr{C}(\varphi) = \mathbf{1}$. Then $e_{\varphi\to\psi}(\mathtt{r}_1) = \mathbf{1}$ and $e_{\varphi}(\mathtt{r}_2) = \mathbf{1}$, for $\mathtt{L}_{\varphi\to\psi}(\mathtt{r}_1) = \varphi \to \psi$, $\mathtt{L}_{\varphi}(\mathtt{r}_2) = \varphi$. $\mathtt{r}_1$ is not a dead end ($\to$ is binary). Since e_{φ} is the fixed point of the sequence of e^{α}_{φ}, the evaluation clause of $e_{\varphi\to\psi}(\mathtt{r}_1)$ is:

$$\begin{aligned}
e_{\varphi\to\psi}(\mathtt{r}_1) = \mathbf{1} \text{ iff } & e_{\varphi\to\psi}(\mathtt{v}_1) = \mathbf{0}, \text{ or}\\
& e_{\varphi\to\psi}(\mathtt{v}_2) = \mathbf{1}, \text{ or}\\
& e_{\varphi\to\psi}(\mathtt{v}_1) = \mathbf{j}, e_{\varphi\to\psi}(\mathtt{v}_2) = \mathbf{k}, \text{ and } \mathbf{j} \le \mathbf{k}
\end{aligned} \tag{1}$$

where $\mathtt{L}_{\varphi\to\psi}(\mathtt{v}_1) = \varphi$ and $\mathtt{L}_{\varphi\to\psi}(\mathtt{v}_2) = \psi$. $\mathtt{v}_1, \mathtt{v}_2 \in \mathtt{N}_{\varphi\to\psi}$ by Definition 4.3 (they are successors of the root node $\mathtt{r}_1$). Since $e_{\varphi}(\mathtt{r}_2) = \mathbf{1}$ and $\mathtt{L}_{\varphi}(\mathtt{r}_2) = \mathtt{L}_{\varphi\to\psi}(\mathtt{v}_1)$, by Proposition 4.23 also $e_{\varphi\to\psi}(\mathtt{v}_1) = \mathbf{1}$. By equation (1) (it's an 'if and only if'), $e_{\varphi\to\psi}(\mathtt{v}_2) = \mathbf{k}$ and $\mathbf{1} \le \mathbf{k}$, but $\mathbf{k} \in \mathsf{V}_{\L}$, so $\mathbf{k} = \mathbf{1}$. Applying again Proposition 4.23 yields that $e_{\psi}(\mathtt{r}_3) = \mathbf{1}$, for $\mathtt{L}_{\psi}(\mathtt{r}_3) = \mathtt{L}_{\varphi\to\psi}(\mathtt{v}_2) = \psi$, therefore $\mathscr{C}(\psi) = \mathbf{1}$ as desired.

Suppose now that $\mathscr{C}(\varphi) = \mathbf{0}$, or $\mathscr{C}(\psi) = \mathbf{1}$, or $\mathscr{C}(\varphi) = \mathbf{j}$ and $\mathscr{C}(\psi) = \mathbf{k}$, for $\mathbf{j} \le \mathbf{k}$. Then, $e_{\varphi}(\mathtt{r}_2) = \mathbf{0}$, or $e_{\psi}(\mathtt{r}_3) = \mathbf{1}$, or $e_{\varphi}(\mathtt{r}_2) = \mathbf{j}$ and $e_{\psi}(\mathtt{r}_3) = \mathbf{k}$ (with labels as above). Let the last case be given (otherwise it is similar). By Proposition 4.23, $e_{\varphi\to\psi}(\mathtt{v}_1) = \mathbf{j}$ and $e_{\varphi\to\psi}(\mathtt{v}_2) = \mathbf{k}$ ($\mathtt{v}_1$ and $\mathtt{v}_2$ exist by Definition 4.3), and by equation (1) $e_{\varphi\to\psi}(\mathtt{r}_1) = \mathbf{1}$, i.e., $\mathscr{C}(\varphi \to \psi) = \mathbf{1}$, as desired. □

Notice that Proposition 4.26 can also be proven essentially along the lines of the above proof.

PROPOSITION 4.30 (Restall (1992)). *There is no continuum-valued Łukasiewicsz evaluation for $\mathcal{L}_{\mathsf{Tr}}$ that*: (*i*) *agrees with an ω-model for $\mathcal{L}$, and* (*ii*) *validates the* T-SCHEMA *or* INTERSUBSTITUTIVITY.

[57] See, for example, item 1.1 in Definition 4.14.

https://doi.org/10.1017/S1755020319000078 Published online by Cambridge University Press

Proof (Based on Field (2008), adapted to the present framework). Let ρ be the fixed point of $\neg\forall n \mathsf{Tr}(\ulcorner \underset{.}{r}(n, x)\urcorner)$ (for the function r, see Definition 4.31). Consider the following graph:

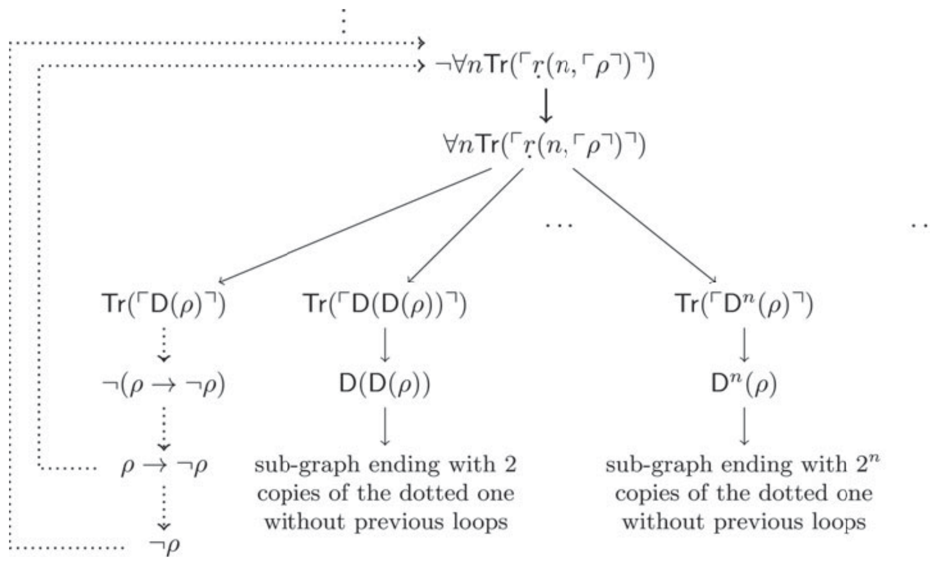

Fig. 9. The semantic graph of ρ.

The evaluation of the nodes in this graph yields an infinite system of equations at level $\omega + 1$. Suppose that this system has exactly one solution, and consider the following cases:

– Suppose that $\mathscr{C}(\rho) = \mathbf{1}$. An easy induction shows that the equation associated with $\mathsf{D}^n(\rho)$, for all $n \in \omega$, has $\mathbf{1}$ as its only solution. Therefore $\mathbf{1} = \mathscr{C}(\rho) = \mathscr{C}(\neg\forall n \mathsf{Tr}(\ulcorner \underset{.}{r}(n, \ulcorner\rho\urcorner)\urcorner)) = \mathbf{1} - \mathscr{C}(\forall n \mathsf{Tr}(\ulcorner \underset{.}{r}(n, \ulcorner\rho\urcorner)\urcorner)) = \mathbf{1} - \mathbf{1} = \mathbf{0}$, which is absurd.
– Suppose that $\mathscr{C}(\rho) = \mathbf{k}$, $\mathbf{k} < \mathbf{1}$. If $\mathbf{0} \leq \mathbf{k} \leq \mathbf{1/2}$, the only solution to the equation associated with each sentence of the form $\mathsf{D}^n(\varphi)$ is $\mathbf{0}$, and so $\mathscr{C}(\forall n \mathsf{Tr}(\ulcorner \underset{.}{r}(n, \ulcorner\rho\urcorner)\urcorner)) = \mathbf{0}$ and $\mathscr{C}(\neg\forall n \mathsf{Tr}(\ulcorner \underset{.}{r}(n, \ulcorner\rho\urcorner)\urcorner)) = \mathbf{1} = \mathscr{C}(\rho)$, against our supposition. If $\mathbf{1/2} < \mathbf{k} < \mathbf{1}$, then there is a $j \in \omega$ s.t. the only possible solution for $\mathsf{D}^j(\rho)$ is less than or equal to $\mathbf{1/2}$. Then consider $\mathsf{D}^{j+1}(\rho)$ and reason as above. □

PROPOSITION 4.32. *For every $\varphi \in \mathcal{L}_{\mathsf{Tr}}$ and every $\mathsf{V}_{\text{Ł}}$, if $\mathscr{C}(\varphi) = \mathbf{k}$, for $\mathbf{k} \in \mathsf{V}_{\text{Ł}}$, then*:

1. *For all ordinals $\alpha \in \mathsf{Ord}_{\mathsf{Nt}}$, $\mathscr{C}(\mathsf{D}^\alpha(\varphi)) \in \mathsf{V}_{\text{Ł}}$. In particular (for γ limit)*:

$$\mathscr{C}(\mathsf{D}^{\alpha+1}(\varphi)) = \mathbf{1} - \min[\mathbf{1}, (\mathbf{1} - \mathscr{C}(\mathsf{D}^\alpha(\varphi)) + \mathbf{1} - \mathscr{C}(\mathsf{D}^\alpha(\varphi)))]$$
$$\mathscr{C}(\mathsf{D}^\gamma(\varphi)) = \inf\{\mathscr{C}(\mathsf{D}^\alpha(\varphi)) | \alpha < \gamma\}$$

2. *There is a unique ordinal $\delta' \in \mathsf{Ord}_{\mathsf{Nt}}$ s.t. for all $\delta \in \mathsf{Ord}_{\mathsf{Nt}}$ greater than or equal to δ'*:

$$\mathscr{C}(\mathsf{D}^\delta(\varphi)) = \begin{cases} \mathbf{1} & \textit{if and only if } \mathscr{C}(\varphi) = \mathbf{1} \\ \mathbf{0} & \textit{if and only if } \mathscr{C}(\varphi) \in \mathsf{V}_{\text{Ł}} \textit{ and } \mathscr{C}(\varphi) < \mathbf{1} \end{cases}$$

Proof. As for the first item, let $\alpha \in \mathsf{Ord}_{\mathsf{Nt}}$ and assume that the claim holds for $\gamma < \alpha$. Now, $\mathsf{D}^{\alpha+1}(\varphi) = \neg(\mathsf{D}^{\alpha}(\varphi) \rightarrow \neg\mathsf{D}^{\alpha}(\varphi))$ and, by IH, $\mathscr{C}(\mathsf{D}^{\alpha}(\varphi)) \in \mathsf{V}_{\text{Ł}}$. By Proposition 4.26, $\mathscr{C}(\neg(\mathsf{D}^{\alpha}(\varphi) \rightarrow \neg\mathsf{D}^{\alpha}(\varphi))) \in \mathsf{V}_{\text{Ł}}$ i.e., $\mathscr{C}(\mathsf{D}^{\alpha+1}(\varphi)) \in \mathsf{V}_{\text{Ł}}$, and the successor case holds by construction. A similar argument establishes the limit case.

As for the second item, note that if $\mathscr{C}(\varphi) \leq \mathbf{1/2}$, then $\mathscr{C}(\mathsf{D}(\varphi)) = \mathbf{0}$ (by Proposition 4.26 and simple calculation). We therefore distinguish two cases:

– $\mathsf{V}_{\text{Ł}} = \{\mathbf{0}, \mathbf{1/2m}, \dots, \mathbf{2m{-}1/2m}, \mathbf{1}\}$. If $\mathscr{C}(\varphi) = \mathbf{k/2m}$ (for $\mathbf{1/2} < \mathbf{k/2m} < \mathbf{1}$), then by Proposition 4.26:

$$\mathscr{C}(\mathsf{D}(\varphi)) = \mathbf{1} - \min[\mathbf{1}, (\mathbf{1} - \mathbf{k/2m} + (\mathbf{1} - \mathbf{k/2m}))] = \mathbf{(2k{-}2m)/2m}. \tag{2}$$

So, if $\mathscr{C}(\varphi) = \mathbf{k/2m}$ (for $\mathbf{1/2} < \mathbf{k/2m} < \mathbf{1}$), then $\mathscr{C}(\mathsf{D}^{m}(\varphi)) = \mathbf{0}$ applying equation (2) at most m times. Therefore, for all $\alpha \in \mathsf{Ord}_{\mathsf{Nt}}$ s.t. $\alpha \geq m$, $\mathscr{C}(\mathsf{D}^{\alpha}(\varphi)) = \mathbf{0}$, by item 1 of the Proposition, equation (2), and Proposition 4.26. This proves item 2 for every finite $\mathsf{V}_{\text{Ł}}$, in which case δ' is the smallest $l \leq m$ s.t. $\mathscr{C}(\mathsf{D}^{l}(\varphi)) = \mathbf{0}$ (uniqueness is immediate).

– $\mathsf{V}_{\text{Ł}} = [\mathbf{0}, \mathbf{1}]$ and $\mathscr{C}(\varphi) = \mathbf{k}$, for $\mathbf{0} \leq \mathbf{k} < \mathbf{1}$. So, there is a $\mathbf{j/2n}$ s.t. $\mathbf{k} < \mathbf{j/2n} < \mathbf{1}$. By the previous result on finite numerical value spaces, $\mathsf{D}^{n}(\varphi) = \mathbf{0}$. Since such a D^{n} exists for each $\mathbf{k} \in \mathsf{V}_{\text{Ł}}$, the claim holds also for their limit D^{ω} (which exists and is unique by item 1 of the Proposition). Therefore, for all $\delta \in \mathsf{Ord}_{\mathsf{Nt}}$ s.t. $\delta \geq \omega$, $\mathscr{C}(\mathsf{D}^{\delta}(\varphi)) = \mathbf{0}$, by item 1, equation (2), and Proposition 4.26, and this shows that if $\mathsf{V}_{\text{Ł}} = [\mathbf{0}, \mathbf{1}]$, then $\delta' = \omega$. □

§7. Acknowledgments. This article has been long in the making. I started working on it during my DPhil (the article was then titled ‘Graphs, truth, and conditional(s)’), but I only had a chance to complete it much later, when my views on truth and semantic paradox had undergone significant changes. I am very grateful to Volker Halbach for invaluable feedback and discussions throughout all the stages of the article. I am also indebted to Hartry Field for many helpful conversations, and to Julien Murzi for numerous useful comments. I owe thanks to Jeff Barrett, Catrin Campbell-Moore, Andrea Cantini, Emmanuel Chemla, Paul Egre, Melvin Fitting, Chris Gauker, Joel Hamkins, Richard Kimberly Heck, Graham Leigh, Hannes Leitgeb, Pierluigi Minari, Carlo Nicolai, Lavinia Picollo, Graham Priest, Chris Scambler, Thomas Schindler, James Studd, and Philip Welch for many helpful exchanges on the topic of this article, and to two anonymous referees for several detailed observations that have led to significant improvements. Finally, I also wish to thank audiences at various conferences and workshops in New York (CUNY), Bristol, Cambridge, Florence, Irvine, Paris (ENS), Bonn (HCM), and Salzburg for their insightful remarks. This research was supported by the Arts and Humanities Research Council (AHRC) and the Scatcherd European Scholarship for research carried out at the University of Oxford and at NYU, and by the *Österreichischer Austauschdienst* (OeAD) and the *Fonds zur Förderung der wissenschaftlichen Forschung* (FWF), grant no. P29716-G24, for research carried out at the University of Salzburg.

BIBLIOGRAPHY

Asenjo, F. G. (1966). A calculus of antinomies. *Notre Dame Journal of Formal Logic*, **7**(1), 103–105.

Barwise, J. & Etchemendy, J. (1987). *The Liar*: *An Essay on Truth and Circularity*. Oxford: Oxford University Press.

Beall, J. (2001). Is Yablo's paradox noncircular? *Analysis*, **61**(3), 176–187.
Beall, J. (2006). True, false and paranormal. *Analysis*, **66**(2), 102–114.
Beall, J. (2007a). Prolegomenon to future revenge. In Beall, J., editor. *Revenge of the Liar*. Oxford: Oxford University Press, pp. 1–30.
Beall, J. (editor) (2007b). *Revenge of the Liar*. Oxford: Oxford University Press.
Beall, J. (2009). *Spandrels of Truth*. Oxford: Oxford University Press.
Beringer, T. & Schindler, T. (2017). A graph-theoretic analysis of the semantic paradoxes. *Bulletin of Symbolic Logic*, **23**(4), 442–492.
Billon, A. (2013). The truth-tellers paradox. *Logique et Analyse*, **56**(224), 371–389.
Blamey, S. (2002). Partial logic. In Gabbay, D. M. and Guenthner, F., editors. *Handbook of Philosophical Logic*, *Second Edition*, Vol. V. Dordrecht: Kluwer Academic Publishers, pp. 261–353.
Bondy, A. & Murty, U. (2008). *Graph Theory*. London: Springer.
Bueno, O. & Colyvan, M. (2003). Paradox without satisfaction. *Analysis*, **62**(2), 152–156.
Burgess, J. (1986). The truth is never simple. *Journal of Symbolic Logic*, **51**(3), 663–681.
Burgess, J. (1988). Addendum to 'The truth is never simple'. *Journal of Symbolic Logic*, **53**(2), 390–392.
Chemla, E. & Égré, P. (2019). Suszko's problem: Mixed consequence and compositionality. *The Review of Symbolic Logic*, to appear.
Chemla, E., Égré, P., & Spector, B. (2017). Characterizing logical consequence in many-valued logic. *Journal of Logic and Computation*, **27**(1), 2193–2226.
Chierchia, G. & McCconnell-Ginet, S. (2000). *Meaning and Grammar*: *Introduction to Semantics* (second edition). Cambridge, MA: MIT Press.
Cieśliński, C. (2007). Deflationism, conservativeness and maximality. *Journal of Philosophical Logic*, **36**(6), 695–705.
Cobreros, P., Égré, P., Ripley, D., & van Rooij, R. (2012). Tolerant, classical, strict. *Journal of Philosophical Logic*, **41**(2), 347–385.
Cobreros, P., Égré, P., Ripley, D., & van Rooij, R. (2013). Reaching transparent truth. *Mind*, **122**(488), 841–866.
Cook, R. (2004). Pattern of paradox. *The Journal of Symbolic Logic*, **69**(3), 767–774.
Cook, R. (2006). There are noncircular paradoxes (but Yablo's isn't one of them!). *The Monist*, **89**(1), 118–149.
Cook, R. (2007). Embracing revenge: On the indefinite extensibility of language. In Beall, J., editor. *Revenge of the Liar*. Oxford: Oxford University Press, pp. 31–52.
Cook, R. (2009). What is a truth value, and how many are there? *Studia Logica*, **92**, 183–201.
Cook, R. (2014). *The Yablo Paradox*: *An Essay on Circularity*. Oxford: Oxford University Press.
Cook, R. & Tourville, N. (2016). Embracing the technicalities: Expressive completeness and revenge. *The Review of Symbolic Logic*, **9**(2), 325–358.
Davidson, D. (1967). Truth and meaning. *Synthese*, **17**, 304–323.
Davis, L. (1979). An alternate formulation of Kripke's theory of truth. *Journal of Philosophical Logic*, **8**(1), 289–296.
Diestel, R. (2010). *Graph Theory* (fourth edition). Berlin: Springer.
Dyrkolbotn, S. & Walicki, M. (2014). Propositional discourse logic. *Synthese*, **191**(5), 863–899.
Eldridge-Smith, P. (2015). Two paradoxes of satisfaction. *Mind*, **124**(493), 85–119.
Field, H. (2002). Saving the truth schema from paradox. *Journal of Philosophical Logic*, **31**(1), 1–27.

Field, H. (2003). A revenge-immune solution to the semantic paradoxes. *Journal of Philosophical Logic*, **32**(2), 139–177.

Field, H. (2007). Solving the paradoxes, escaping revenge. In Beall, J., editor. *Revenge of the Liar*. Oxford: Oxford University Press, pp. 53–144.

Field, H. (2008). *Saving Truth from Paradox*. Oxford: Oxford University Press.

Field, H. (2013). Naive truth and restricted quantification: Saving truth a whole lot better. *The Review of Symbolic Logic*, **7**(1), 147–191.

Fine, K. (2017). Truthmaker semantics. In Hale, B., Wright, C., and Miller, A., editors. *A Companion to the Philosophy of Language, Second Edition,* Oxford: Wiley-Blackwell, pp. 556–577.

Fischer, M., Halbach, V., Kriener, J., & Stern, J. (2015). Axiomatizing semantic theories of truth? *The Review of Symbolic Logic*, **8**(2), 257–278.

Friedman, H. & Sheard, M. (1987). An axiomatic approach to self-referential truth. *Annals of Pure and Applied Logic*, **33**, 1–21.

Gaifman, H. (1988). Operational pointer semantics: Solution to self-referential puzzles I. In Vardi, M., editor. *Theoretical Aspects of Reasoning about Knowledge*. Los Angeles: Morgan Kauffman, pp. 43–59.

Gaifman, H. (1992). Pointers to truth. *Journal of Philosophy*, **89**(5), 223–261.

Gaifman, H. (2000). Pointers to propositions. In Chapuis, A. and Gupta, A., editors. *Circularity, Definition, and Truth*. Indian Council of Philosophical Research.

Gottwald, S. (2001). *A Treatise on Many-valued Logics*. Studies in Logic and Computation. Baldock, Hertfordshire, England: Research Studies Press LTD.

Greenough, P. (2011). Truthmaker gaps and the no-no paradox. *Philosophy and Phenomenological Research*, **82**(3), 547–563.

Gupta, A. (1982). Truth and paradox. *Journal of Philosophical Logic*, **11**(1), 1–60.

Gupta, A. & Belnap, N. (1993). *The Revision Theory of Truth*. Cambridge (MA): MIT Press.

Hájek, P., Paris, J., & Shepherdson, J. (2000). The liar paradox and fuzzy logic. *Journal of Symbolic Logic*, **65**(1), 339–346.

Halbach, V. (2011). *Axiomatic Theories of Truth*. Cambridge: Cambrdige University Press.

Halbach, V. & Horsten, L. (2006). Axiomatizing Kripke's theory of truth. *Journal of Symbolic Logic*, **71**(2), 677–712.

Halbach, V., Leitgeb, H., & Welch, P. (2003). Possible-worlds semantics for modal notions conceived as predicates. *Journal of Philosophical Logic*, **32**(2), 179–223.

Halbach, V. & Visser, A. (2014a). Self-reference in arithmetic I. *The Review of Symbolic Logic*, **7**(4), 671–691.

Halbach, V. & Visser, A. (2014b). Self-reference in arithmetic II. *The Review of Symbolic Logic*, **7**(4), 692–712.

Halbach, V. & Zhang, S. (2017). Yablo without Gödel. *Analysis*, **77**(1), 53–59.

Hansen, C. (2015). Supervaluation on trees for Kripke's theory of truth. *The Review of Symbolic Logic*, **8**(1), 46–74.

Hazen, A. (1981). Davis's formulation of Kripke's theory of truth: A correction. *Journal of Philosophical Logic*, **10**(3), 309–311.

Herzberger, H. (1982a). Naive semantics and the liar paradox. *Journal of Philosophy*, **79**(9), 479–497.

Herzberger, H. (1982b). Notes on naive semantics. *Journal of Philosophical Logic*, **11**(1), 61–102.

Horsten, L. (2009). Levity. *Mind*, **118**(471), 555–581.

Horsten, L. (2012). *The Tarskian Turn. Deflationism and Axiomatic Truth*. Cambridge (MA): MIT Press.

Ketland, J. (2003). Can a many-valued language functionally represent its own semantics? *Analysis*, **63**(4), 292–297.

Ketland, J. (2004). Bueno and Colyvan on Yablo's paradox. *Analysis*, **64**(2), 165–172.

Ketland, J. (2005). Yablo's paradox and ω-inconsistency. *Synthese*, **145**(3), 295–302.

Kleene, S. C. (1952). *Introduction to Metamathematics*. New York: van Nostrand.

Kremer, P. (2009). Comparing fixed-point and revision theories of truth. *Journal of Philosophical Logic*, **38**(4), 363–403.

Kripke, S. (1975). Outline of a theory of truth. *Journal of Philosophy*, **72**(19), 690–716.

Leitgeb, H. (2002). What is a self-referential sentence? Critical remarks on the alleged (non)circularity of Yablo's paradox. *Logique et Analyse*, **177–178**, 3–14.

Leitgeb, H. (2005). What truth depends on. *Journal of Philosophical Logic*, **34**(2), 155–192.

Leitgeb, H. (2007). On the metatheory of Field's 'Solving the paradoxes, escaping revenge'. In Beall, J., editor. *Revenge of the Liar*. Oxford: Oxford University Press, pp. 159–183.

Maudlin, T. (2004). *Truth and Paradox*: *Solving the Riddles*. New York: Oxford University Press.

Maudlin, T. (2007). Reducing revenge to discomfort. In Beall, J., editor. *Revenge of the Liar*. Oxford: Oxford University Press, pp. 184–196.

McCarthy, T. (1988). Ungroundedness in classical languages. *Journal of Philosophical Logic*, **17**(1), 61–74.

McGee, V. (1985). How truthlike can a predicate be? A negative result. *Journal of Philosophical Logic*, **14**(4), 399–410.

McGee, V. (1991). *Truth, Vagueness, and Paradox*. Indianapolis: Hackett Publishing Company.

Montague, R. (1974). *Formal Philosophy. Selected Papers of Richard Montague*. New Haven: Yale University Press.

Mortensen, C. & Priest, G. (1981). The truth teller paradox. *Logique et Analyse*, **95–96**, 381–388.

Moschovakis, Y. (1974). *Elementary Induction on Abstract Structures*. Amsterdam, London and New York: North-Holland and Elsevier.

Murzi, J. & Rossi, L. (2019). Generalised revenge. *Australasian Journal of Philosophy*, to appear.

Nicolai, C. & Rossi, L. (2018). Principles for object-linguistic consequence: From logical to irreflexive. *Journal of Philosophical Logic*, **47**(3), 549–577.

Priest, G. (1979). The logic of paradox. *Journal of Philosophical Logic*, **8**(1), 219–241.

Priest, G. (1997). Yablo's paradox. *Analysis*, **57**(4), 236–242.

Priest, G. (2006). *In Contradiction* (expanded edition). Oxford: Oxford University Press.

Priest, G. (2007). Revenge, Field, and ZF. In Beall, J., editor. *Revenge of the Liar*. Oxford: Oxford University Press, pp. 225–233.

Rabern, L., Rabern, B., & Macauley, M. (2013). Dangerous reference graphs and semantic paradoxes. *Journal of Philosophical Logic*, **42**(5), 727–765.

Restall, G. (1992). Arithmetic and truth in Lukasiewicz's infinitely valued logic. *Logique et Analyse*, **139–140**, 303–312.

Restall, G. (2007). Curry's revenge: The costs of nonclassical solutions to the paradoxes of self-reference. In Beall, J., editor. *Revenge of the Liar*. Oxford: Oxford University Press, pp. 262–271.

Rogers, H. (1987). *Theory of Recursive Functions and Effective Computability* (second edition). Cambridge (MA) and London: MIT Press.
Rossi, L. (2016). Adding a conditional to Kripke's theory of trugh. *Journal of Philosophical Logic*, **45**(5), 485–529.
Rossi, L. (2019). Model-theoretic semantics and revenge paradoxes. *Philosophical Studies*, to appear.
Scharp, K. (2007). Aletheic vengeance. In Beall, J., editor. *Revenge of the Liar*. Oxford: Oxford University Press, pp. 272–319.
Scharp, K. (2013). *Replacing Truth*. Oxford: Oxford University Press.
Schlenker, P. (2007). The elimination of self-reference: Generalized Yablo-series and the theory of truth. *Journal of Philosophical Logic*, **36**(3), 251–307.
Schlenker, P. (2010). Super-liars. *The Review of Symbolic Logic*, **3**(3), 374–414.
Shapiro, L. (2011). Expressibility and the liar's revenge. *Australasian Journal of Philosophy*, **89**(2), 1–18.
Simmons, K. (2007). Revenge and context. In Beall, J., editor. *Revenge of the Liar*. Oxford: Oxford University Press, pp. 345–367.
Smith, J. (1984). A simple solution to Mortensen and Priest's truth teller paradox. *Logique et Analyse*, **27**(106), 217–220.
Sorensen, R. (2001). *Vagueness and Contradiction*. Oxford: Oxford University Press.
Urquhart, A. (2001). Basic many-valued logic. In Gabbay, D. and Guenthner, F., editors. *Handbook of Philosophical Logic*, *Second Edition*, Vol. 2. Dordrecht: Kluwer Academic Publishers, pp. 249–295.
Visser, A. (1984). Four valued semantics and the liar. *Journal of Philosophical Logic*, **13**(2), 181–212.
Visser, A. (1989). Semantics and the liar paradox. In Gabbay, D. and Günthner, F., editors. *Handbook of Philosophical Logic*, Vol. 4. Dordrecht: Reidel, pp. 617–706.
Walicki, M. (2009). Reference, paradoxes and truth. *Synthese*, **171**, 195–226.
Walicki, M. (2017). Resolving infinitary paradoxes. *Journal of Symbolic Logic*, **82**(2), 709–723.
Wen, L. (2001). Semantic paradoxes as equations. *The Mathematical Intelligencer*, **23**(1), 43–48.
Yablo, S. (1982). Grounding, dependence, and paradox. *Journal of Philosophical Logic*, **11**, 117–137.
Yablo, S. (1985). Truth and reflection. *Journal of Philosophical Logic*, **14**(3), 297–349.
Yablo, S. (1993). Paradox without self-reference. *Analysis*, **53**(4), 251–252.
Yablo, S. (2006). Circularity and paradox. In Bolander, T., Hendricks, V., and Pedersen, S., editors. *Self-Reference*. Stanford: CSLI Publications, pp. 139–157.
Yi, B. (1999). Descending chains and the contextualist approach to paradoxes. *Notre Dame Journal of Formal Logic*, **40**(4), 554–567.
Zardini, E. (2011). Truth without contra(di)ction. *Review of Symbolic Logic*, **4**(4), 498–535.

DEPARTMENT OF PHILOSOPHY (KGW)
UNIVERSITY OF SALZBURG
FRANZISKANERGASSE 1
5020 SALZBURG, AUSTRIA
E-mail: lorenzo.rossi@sbg.ac.at