The Journal of Symbolic Logic
Volume 00, Number 0, XXX 0000

THE DEFINABILITY OF TRUTH IN A LOGIC OF SENTENTIAL OPERATORS

MICHAL WALICKI

Abstract. Logic of Sentential Operators (LSO) extends classical logic with sentential quantifiers and operators, making self-reference and paradoxes expressible. All classical tautologies and contradictions remain intact, while the reasoning system preserves the full rule set of the classical sequent calculus LK. We show how to introduce an operator that defines the semantic structure and a nontrivial truth operator. The extension with these operators is conservative, validates the unrestricted Convention T and the compositionality axioms for truth. Its definition relies on the operator of syntactic equality of sentences. A proof of its consistency is provided, thereby filling a gap in earlier presentations of LSO.

§1. Introduction. Truth theories study the consistency of adding a truth predicate T to classical logic. The standard setting combines (a sufficiently strong) arithmetic, **a**rithmetized **s**yntax, and a **t**ruth predicate, so that terms denote both their semantic referents and, via Gödelisation, formulas themselves. (By "AST" we refer to such systems, in general, where the metalanguage is represented by the terms of the object-language and truth by a predicate.) Ideally, predicate T would satisfy Convention T, namely, [T] $T(\ulcorner\phi\urcorner) \leftrightarrow \phi$, for each sentence ϕ and (the numeral of) its Gödel number $\ulcorner\phi\urcorner$. It reflects the intended reading of term $\ulcorner\phi\urcorner$ as a name of sentence ϕ. However, this is impossible by Tarski's undefinability theorem, [9]. Even approaching this ideal in any systematic and monotone way is doomed to failure as there are no formal optimality criteria for the incomparable extensions of such predicates [3]. "As everyone knows, this is an area in which one simply cannot have all that one might reasonably ask for" [7].

Two responses are possible: (a) continue working with AST, restricting [T] to prevent paradoxes and accepting the impossibility it embodies, or (b) search for another framework capable of defining its truth. The undefinability theorem holds for classical logic under some weak requirements on the language, so increasing its expressivity is of no help. Consequently, choice (b) effectively renounces classical logic – an unattractive move, usually accompanied by efforts to preserve as much of it as possible. Choice (a), on the other hand, even if technically fruitful, is not philosophically very satisfying. Formal logic, at least in its philosophical dimension, tries to capture some aspects of natural reasoning. It is hard to find there any counterparts of the restrictions required for maintaining the consistency of AST's truth theories. One is led to suppose that natural language, "for which the normal laws of logic hold, must be inconsistent" or that it cannot capture its truth concept. [9, p.165]. The awkwardness of this choice only aggravates with the observation that it rests solely on the undefinability theorem – or rather, the assumption that it actually applies to natural reasoning. The reservations that the notion of truth is restricted to formalized languages are one thing; the multiple suggestions, like the quoted one, of the inconsistency of natural languages arising from the theorem – another. Yet the intricacy of the truth concept does not visibly imply its indefinability, and the inconsistency of natural languages with the normal laws of logic, though conceivable, seems no more plausible than their consistency. "[S]omehow, it seems, natural languages defy the indefinability theorem", [6], and whether it at all applies to them is not indisputable.

This unclear, if not directly dubious, import of truth formalisations in AST, and the impossiblity they are bound to accept, makes alternative (b) attractive to many. This paper, however, does

0022-4812/00/0000-0000/$00.00

not follow it. Our worry about AST's model concerns not the classical logic but the status of self-reference. It is namely a feature of the natural language itself, not of any theory imposed on it. A framework relying solely on the naming or coding of sentences for metalinguistic reference risks overlooking possibly relevant features. A Cretan's remark "All Cretans always lie" involves a kind of *suppositio materialis*, a reference to the sentence itself rather than only to its truth-value, but as an intensional rather than a merely syntactic phenomenon.[1] The saying refers, among other things, to the sentence without naming it. That predicating truth of sentences via their names may be plausible (and technically convenient) does not establish it as the primary, let alone the only, mode of metalinguistic reference, whether in natural reasoning or in formal systems. This is particularly significant, considerting that Tarski established not the undefinability of truth, but of the truth *predicate*.

A different model of self-reference arises in [12] with a logic of sentential operators, LSO. Statements about statements are made using such operators, instead of predicates on sentence names.[2] A sentence in a *sentential position* (not in the scope of any operator) stands for its truth-value, while one occurring in a *nominal position*, as an argument of an operator, may be taken in a material supposition. For such an operator K, the first occurrence of S in $K(S) \to S$ is nominal, and the second sentential. One can quantify over sentences occuring in both ways, e.g., $\forall\phi(K(\phi) \to \neg\phi)$. Reading $K(\phi)$ as 'Cretan K saying ϕ', this stands for 'everything Cretan K says is false', and $K(\forall\phi(K(\phi) \to \neg\phi))$ for Cretan K saying that.

Paradoxes in LSO are, as in natural discourse, specific metalinguistic claims, that may seem coherent to some point, after which they unveil contradiction. They require valuations of operators that do not extend to the whole language, but their expressibility does not require any artificial restrictions for maintaining consistency. The liar paradox, arising when K says only $\forall\phi(K(\phi) \to \neg\phi)$, is resolved if K says also some true sentence. Like the (in)consistency of a theory is the reponsibility of its designer, paradoxes depend on what we say and are not excluded by the language or the formalism alone. To this earlier treatment of paradoxes this paper adds a formalisation of LSO's truth in LSO. The change from the predicate to the operator setting might suggest an instance of (b) above, but LSO avoids the choice taken as either (a) classical logic or (b) an adequate internalisation of truth, that arises from modelling truth as a predicate. LSO is a classical logic, in the sense of being two-valued and retaining all rules of the classical sequent system LK, augmenting it only with two rules for the sentential quantifiers. The semantics is formulated differently, to cater for the self-reference present in the language but, when restricted to the first-order sublanguage, is equivalent to its classical semantics.

In the operator setting, Convention T – formulated as (T) $\forall\phi(\mathsf{T}\phi \leftrightarrow \phi)$ – can be accepted in the way proposed by Tarski: as the necessary condition of material adequacy for a definition of truth. Taken as a definition it introduces only a trivial identity-like operator, while a satisfactory definition should have some substance, even if what exactly this means is unclear.[3] We may expect it to say something more – to imply (T) without being equivalent to it. This, too, can be achieved cheaply by strengthening (T) with some irrelevant non-tautological conjunct. A satisfactory definition of truth should reflect how the semantic structure makes some sentences true and others false. The paper does that without, however, offering any sufficient criteria. It keeps thus the door open to various formalisations without asserting any plurality of truths. Taking (T) as a necessary condition, any consistent theory that implies (T) might qualify as a candidate definition of truth – but its ultimate value must be judged on broader grounds. Hopefully the reader will find the proposed definition to be a significant improvement over bare (T). Whether

[1] Material supposition is not our topic, so let us only note that, according to the late Scholastics, it occurs when "a term stands for something it in no way signifies" [8, p.198], or "a word stands for itself or for another word with a non-inferior extent [e.g., a synonym]" [1, p.2]. Names of sentences or terms can realize it, but so can quotations, descriptive references or quantification, like that used above by the Cretan Epimenides.

[2] The formal difference between predicate and operator need not reflect any specific linguistic difference. The truth operator may be read as saying "it is true that...", but the possible correspondence to the expressions of informal language need not be tight. To the extent the considered formalisms reflect natural language, they do so in its logical and semantic aspects rather than the grammatical or syntactic ones.

[3] Deflationists might disagree. Our notion is conservative, but we refrain from discussing here whether this complies with deflationism or goes beyond it.

other alternatives might outperform it must await either new proposals or the formulation of adequate sufficient conditions.

Starting with any classical language $\mathcal{L}^-$ (first-order is used as the natural example), we extend it to LSO language $\mathcal{L}^{\doteq}$, with the sentential quantifiers and the operator $\doteq$ holding about syntactically identical sententences. A digraph $G_M(\mathcal{L}^{\doteq})$ provides the interpretation of $\mathcal{L}^{\doteq}$ which, restricted to $\mathcal{L}^-$, coincides with its classical interpretation in FOL domain M. Truth is then defined in LSO for the LSO sentences in an arbitrary language $\mathcal{L}$ extending $\mathcal{L}^{\doteq}$. The definition is relative to $\mathcal{L}$, hence schematic, and infinite with schemas for infinitely many formulas. It is not restricted to any specific structure but defines truth uniformly in arbitrary structures for $\mathcal{L}$.

Section 2 introduces the necessary background of LSO from [11, 12]. Our definition of truth requires only one operator, the syntactic identity of sentences. Its earlier usage assumed its consistency, which is shown here for the first time. The main contribution is in Section 3 and Section 4 concludes with some philosophical remarks.

§2. The background: LSO. In addition to the usual elements of a classical (here, first-order) language $\mathcal{L}^-$ – object variables $o\mathcal{V}$ (typically, x, y), function and constant symbols $\mathcal{F}$, predicate symbols $\mathcal{P}$ – an LSO language $\mathcal{L}$ has also sentential operator symbols $\mathcal{O}$ and sentential variables $s\mathcal{V}$ (typically, ϕ, ψ), which can be quantified. The language is given by the following grammar in BNF. (Single arguments stand for arbitrary arities; $(\boldsymbol{T}, \boldsymbol{F})$ abbreviates all combinations of term and formula arguments, including empty one for sentential constants; $\mathcal{V} = o\mathcal{V} \cup s\mathcal{V}$.)

$\boldsymbol{T}$	::=	$o\mathcal{V} \mid \mathcal{F}(\boldsymbol{T})$	– terms
$\boldsymbol{A}$	::=	$\mathcal{P}(\boldsymbol{T}) \mid \mathcal{O}(\boldsymbol{T}, \boldsymbol{F})$	– atomic formulas
$\boldsymbol{F}$	::=	$\boldsymbol{A} \mid s\mathcal{V} \mid \boldsymbol{F} \wedge \boldsymbol{F} \mid \neg \boldsymbol{F} \mid \forall \mathcal{V}.\boldsymbol{F}$	– all formulas.

We use only $\neg$ and $\wedge$ – the other connectives are definable in the classical way. Among the operators $\mathcal{O}$, we include the binary infix operator of *s-equality*, $\doteq$, with the intended interpretation as the syntactic identity of sentences. (Referring to $\doteq$ we always assume this interpretation.) A *trivial* equality is $F \doteq F$, for any formula F. Operators bind stronger than logical connectives, e.g., $\phi \doteq P(\psi) \to A$ stands for $(\phi \doteq P(\psi)) \to A$ and not for $\phi \doteq (P(\psi) \to A)$, but for instance $\psi \doteq \forall \phi A \phi$ is read as $\psi \doteq (\forall \phi A \phi)$ and not as $(\psi \doteq \forall \phi) A \phi$.

Sentential variables $s\mathcal{V}$ and sentential quantifiers are interpreted substitutionally with the unrestricted substitution class including all sentences of the language. Thus, in any sentence such a quantifier ranges also over this very sentence.

The operators are sentential but, syntactically, can have arguments that are open formulas within the scope of outer quantifiers binding their free variables. If $F(x, \phi)$ has only x and ϕ free, then $O(F(x, \phi))$ is an open formula with these variables free, while $\exists \phi \forall x O(F(x, \phi))$ is a sentence with a legal application of operator O. An evaluation of this sentence applies O at points where variables ϕ and x are assigned specific values.

An application of an operator to any argument(s) is an atomic expression. Hence, arbitrary valuations of such closed atoms are admissible, e.g., $O(S)$ and $O(\neg S)$, for a sentence S, may obtain all four possible combinations of truth values. Specific interactions of the operators with their arguments are left to the appropriate axiomatisations.

We view an operator application $O(S)$ as statement O about sentence S. Likewise, $O(O(S))$ is statement O about $O(S)$. Applications of operators, together with the sentential quantifiers, form thus the metalanguage – not only for the underlying object-language $\mathcal{L}^-$, but for $\mathcal{L}$ itself.

Sentential quantifiers suggest that sentences (formulas without any free object or sentential variables) are among the objects of the interpretation domain, but this happens in a special way involving an additional structure. By $\boldsymbol{S}$ we denote the set of all sentences of a language $\mathcal{L}$, by $\mathcal{L}_M$ the language $\mathcal{L}$ expanded with constant symbols M, while by $\boldsymbol{T}_M, \boldsymbol{A}_M, \boldsymbol{S}_M$ – the free algebras of the respective syntactic categories over the elements of a set M. Typically, M is a *domain*, namely, a nonempty set with an interepretation of the $\mathcal{L}^-$ function symbols, but not of the predicate or operator symbols. A *valuation of variables*, in any actual domain M, is an assignment $\alpha \in M^{o\mathcal{V}} \times (\boldsymbol{S}_M)^{s\mathcal{V}}$ of the elements of M to o-variables and of sentences $\boldsymbol{S}_M$ to s-variables. Each domain M determines a digraph with sentences $\boldsymbol{S}_M$ as vertices.

Definition 2.1. *The* language graph $G_M(\mathcal{L})$*, for a language* $\mathcal{L}$ *and domain* M*, has sentences* $\boldsymbol{S}_M$ *as vertices, and edges* $\boldsymbol{E}$ *given by:*

1. *Each atomic sentence* $A \in \boldsymbol{A}_M$ *has a 2-cycle to its negation:* $A \leftrightarrows \neg A$.

Each nonatomic sentence $S \in \boldsymbol{S}_M$ *is the root of the subgraph* $G_M(S)$*:*

	root S	*with edges*	*to* :
2.	$\neg F$	$\longrightarrow$	F,
3.	$F_1 \wedge F_2$	$\longrightarrow$	$\neg F_i$, *for* $i \in \{1,2\}$,
4*a*.	$\forall x F x$	$\longrightarrow$	$\neg F(m)$, *for each* $m \in M$,
4*b*.	$\forall \phi F \phi$	$\longrightarrow$	$\neg F(T)$, *for each* $T \in \boldsymbol{S}_M$.

In 4b, all sentences $\boldsymbol{S}_M$ instantiate the s-quantifier, not only sentences $\boldsymbol{S}$ of $\mathcal{L}$. By $\mathcal{LG}r(\mathcal{L})$ we denote all language graphs for a language $\mathcal{L}$. For each $\mathcal{L}^-$ sentence S its subgraph $G(S)$, rooted in S, is essentially a tree except that instead of the leaves it has 2-cycles with the literals. More complex (and problematic) cycles arise from the s-quantification as illustrated further down.

In any such graph $G = (\boldsymbol{V}, \boldsymbol{E})$, vertices/sentences are assigned truth values following the rule

$$\forall x \in \boldsymbol{V} : \alpha(x) = \mathbf{1} \Leftrightarrow \forall y(\boldsymbol{E}(x,y) \to \alpha(y) = \mathbf{0}). \tag{2.2}$$

For $Y \subseteq \boldsymbol{V}$, let $\boldsymbol{E}^-(Y) = \{x \in \boldsymbol{V} \mid \exists y \in Y : \boldsymbol{E}(x,y)\}$ denote the set of vertices with an edge to some $y \in Y$. Vertices assigned $\mathbf{1}$ in (2.2) form a *kernel* of G, namely, a subset $K \subseteq \boldsymbol{V}$, that is *independent*, $\boldsymbol{E}^-(K) \subseteq \boldsymbol{V} \setminus K$ (no edges between vertices in K; equivalently, $\boldsymbol{E}(K) \subseteq \boldsymbol{V} \setminus K$), and *absorbing*, $\boldsymbol{E}^-(K) \supseteq \boldsymbol{V} \setminus K$ (each vetex outside of K has an edge to K). Conversely, for any kernel K, valuation $\kappa(x) = \mathbf{1} \Leftrightarrow x \in K$ satisfies (2.2). So we identify the two and denote by $Ker(G)$ all kernels (valuations (2.2)) of a graph G.

A kernel K of a language graph G, viewed as the set of true sentences, restricted to atoms determines their valuation κ. The object-level sentences in K are then exactly those that are true in $\mathcal{L}^-$ structure M under κ. The semantics defined by kernels of language graphs, when restricted to the object-language $\mathcal{L}^-$, coincides with its classical semantics. It is now a part of the semantics for the whole language $\mathcal{L}$, evaluating all its sentences.

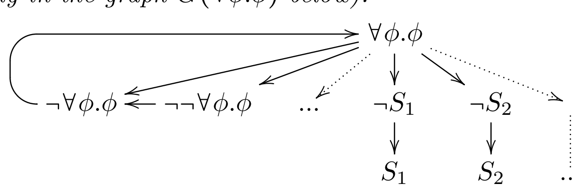
An $\mathcal{L}$ *structure* is a pair (G, K) with $G \in \mathcal{LG}r(\mathcal{L})$ and $K \in Ker(G)$. A *kernel model* of a theory $\Gamma \subseteq \boldsymbol{S}$ is an $\mathcal{L}$ structure with a kernel containing all sentences from Γ. This is the first, definitional equivalence below, the second one reflecting (2.2).

$$(G, K) \vDash_c S \Leftrightarrow S \in K \Leftrightarrow \forall \theta(\boldsymbol{E}(S, \theta) \to \theta \notin K). \tag{2.3}$$

A sentence S is a *logical consequence of a set* Γ *of sentences*, $\Gamma \vDash_c S$, if every kernel model of Γ is also a model of S. The intended intepretation of $\doteq$ requires that kernels contain the diagonal $\{A \doteq A \mid A \in \boldsymbol{S}_M\}$, and the negation of its complement $\{A \not\doteq B \mid \text{distinct } A, B \in \boldsymbol{S}_M\}$.

Every $\mathcal{L}^-$ structure M gives rise to some $\mathcal{L}$ structures (G, K) satisfying the same $\mathcal{L}^-$ sentences. The extension of $\mathcal{L}^-$ to $\mathcal{L}$ is conservative: all classical tautologies (contradictions) remain tautologies (contradictions). Their sentential instances are expressible by single formulas, e.g., $\forall \phi(\phi \vee \neg \phi)$ is a tautology. The richer language introduces also some new tautologies and contradictions.

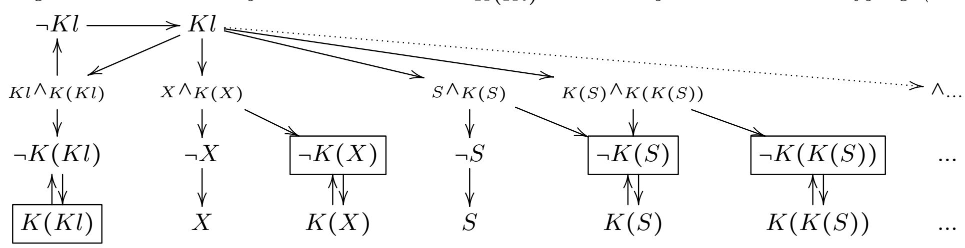
Example 2.4. *Let* $S_1, S_2, \ldots$ *stand for all sentences* $\boldsymbol{S}$*, except the iterated negations of* $\forall \phi.\phi$*, (some appearing explicitly in the graph* $G(\forall \phi.\phi)$ *below):*



The simplified drawing indicates only the essential aspects, ignoring other edges and cycles. Any $S_i \in \boldsymbol{S}$ *valuated to* $\mathbf{0}$ *yields* $\neg S_i = \mathbf{1}$ *and* $\forall \phi.\phi = \mathbf{0}$*, but even if all* $S_i = \mathbf{1}$*, merely the indicated cycles force* $\forall \phi.\phi = \mathbf{0}$*. To obtain a kernel, the odd cycle via* $\neg\neg\forall\phi.\phi$ *must be broken, i.e., some of its vertices must have an out-neighbour* $= \mathbf{1}$*. If all* $\neg S_i = \mathbf{0}$*, this still happens when* $\neg\forall\phi.\phi = \mathbf{1}$*, making* $\neg\neg\forall\phi.\phi = \mathbf{0} = \forall\phi.\phi$*. In a sense, contradiction* $\forall\phi.\phi$ *is a counterexample to its own satisfiability.*

The following example illustrates the way paradoxes arise in LSO.

EXAMPLE 2.5. *Liar K says 'Everything I'm saying is false', $K(Kl)$ with $Kl = \forall\phi(K\phi \to \neg\phi)$, and says only that, $\forall\phi(K\phi \to \phi \doteq Kl)$. $K!Kl$ abbreviates these two sentences. A model must satisfy atoms framed on the drawing: $K(Kl)$ and $\neg K(S)$, for all $S \not\doteq Kl$. It has no extension to a kernel: $K(Kl) = \mathbf{1}$ makes $\neg K(Kl) = \mathbf{0}$, while $\neg K(S) = \mathbf{1}$ makes ${}_{S}{\wedge}_{K(S)} = \mathbf{0}$, for $S \not\doteq Kl$. The resulting unresolved odd cycle $\neg Kl - Kl - {}_{Kl}{\wedge}_{K(Kl)}$ blocks any evaluation satisfying (2.2).*



Unlike contradictions, evaluating always to false, a paradox makes evaluation of some sentences impossible. The unresolved odd cycle reflects the entailment from K's paradoxical claim to K both lying and not lying: $K!Kl \vDash_c Kl \wedge \neg Kl$. This paradox disappears if $K(X)$ holds for some true sentence X, since then ${}_{X}{\wedge}_{K(X)} = \mathbf{1}$ makes $Kl = \mathbf{0}$.

The operator K, under the interpretation satisfying $K!Kl$, leads to the paradox. This is enabled by the central trait of LSO, the self-referential capacity of s-quantification, not requiring any additional theory. While all sentences of the object language $\mathcal{L}^-$ receive unique values under every valuation of atoms, certain valuations of operators may lead to paradoxes that prevent evaluation of some sentences. In principle, it might happen that the language $\mathcal{L}$ itself is inconsistent in this sense, disabling evaluation of some sentences. The existence of kernels in language graphs is a nontrivial claim, demonstrating that this is not the case – every LSO language $\mathcal{L}$, with the rules of classical logic, is consistent, that is, has an interpretation not leading to any paradoxes.

THEOREM 2.6 ([12, Th.3.4]). *For every LSO language $\mathcal{L}$, every graph $G_M(\mathcal{L})$ has a kernel.*

A consistent theory Γ over $\mathcal{L}$ can be extended to a consistent theory by a definitional extension. It introduces (or picks) a predicate or operator P undefined in Γ and defines it by a sentence

$$\forall\phi, x(P(\phi, x) \leftrightarrow \exists\!\!\!/\psi\exists\!\!\!/ y(F(\psi, y, \phi', x'))), \tag{2.7}$$

where $\exists\!\!\!/$ are any quantifiers, while $\exists\!\!\!/\psi\exists\!\!\!/ y(F(\psi, y, \phi', x'))$ is an $\mathcal{L}$ formula with the free variables ϕ', x' contained among free ϕ, x of the left side. A sequence of such extensions is also definitional.

THEOREM 2.8 ([12, Th.5.6]). *For every LSO theory $\Gamma \subset \mathcal{L}$ and its definitional extension D, every kernel model of Γ can be extended to a kernel model of $\Gamma \cup D$.*

Each theorem above ensures freedom from paradoxes of every uninterpreted language $\mathcal{L}$. This remains so when the language addresses its own semantics and truth as described in Section 3.

Reasoning system LSO, originating from [11], extends the classical sequent system LK with two rules for s-quantifiers. By $\Gamma \Rightarrow \Delta$ we denote a sequent with formula sets Γ, Δ, and by $\Gamma \vdash \Delta$ mark its provability.

(Ax) $\Gamma \vdash \Delta$ for Γ and Δ sharing some formula (not necessarily atomic)

$(\neg_L)\ \dfrac{\Gamma \vdash \Delta, A}{\neg A, \Gamma \vdash \Delta}$ $\qquad (\neg_R)\ \dfrac{A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta, \neg A}$

$(\wedge_L)\ \dfrac{A, B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta}$ $\qquad (\wedge_R)\ \dfrac{\Gamma \vdash \Delta, A \quad \Gamma' \vdash \Delta', B}{\Gamma, \Gamma' \vdash \Delta, \Delta', A \wedge B}$

$(\forall_L)\ \dfrac{F(t), \Gamma \vdash \Delta}{\forall x F(x), \Gamma \vdash \Delta}$ legal $F[x\backslash t]$ $\qquad (\forall_R)\ \dfrac{\Gamma \vdash \Delta, F(y)}{\Gamma \vdash \Delta, \forall x F(x)}$ fresh y

$(\forall_L^\phi)\ \dfrac{F(S), \Gamma \vdash \Delta}{\forall\phi F(\phi), \Gamma \vdash \Delta}$ legal $F[\phi\backslash S]$ $\qquad (\forall_R^\phi)\ \dfrac{\Gamma \vdash \Delta, F(\psi)}{\Gamma \vdash \Delta, \forall\phi F(\phi)}$ fresh ψ

The legality of substitutions in $(\forall_L^\phi)$, $(\forall_L)$ involves the usual avoidance of variable capture. Substitution of a sentence in $(\forall_L^\phi)$ is always legal, but also open formulas can be substituted. For instance, $\forall x\exists\phi(Px \leftrightarrow \phi)$ is derivable: instantiating x with a fresh y, we can then instantiate ϕ

with Py. On the other hand, $\exists\phi\forall x(Px \leftrightarrow \phi)$ is not derivable (unless Px has a provably constant value for all x) because substitution of Px for ϕ into $\forall x(Px \leftrightarrow \phi)$ is illegal.

If needed, we can use infinite Γ, Δ, since any proof, also from an infinite Γ, involves only its finite subset. Rules $(\forall_L^\phi)$, $(\forall_R^\phi)$ are like second-order, but the logic is compact, because quantifiers range only over sentences, not over open formulas (predicate positions) or arbitrary subsets. In spite of this quantification over its own syntax, LSO is not even a weak second-order logic that quantifies, e.g., only over definable sets. Its closest relative is rather two-sorted first-order logic.

LSO is sound and complete but for the semantics of semikernels that generalizes that of kernels. Every kernel is also a *semikernel* [5], namely, a subset L of vertices such that

$$\boldsymbol{E}(L) \subseteq \boldsymbol{E}^-(L) \subseteq \boldsymbol{V} \setminus L. \tag{2.9}$$

A semikernel L, also called *local kerenel*, is independent but restricts the kernel condition, requiring all vertices in $\boldsymbol{V} \setminus L$ to have edges back to L, to vertices $\boldsymbol{E}(L)$, i.e., those reached from L. It retains the valuation condition (2.2) for vertices in $\boldsymbol{E}^-[L] = L \cup \boldsymbol{E}^-(L)$ and this determines the classical character of the logic. Semikernel L *covers* vertices $\boldsymbol{E}^-[L]$, and it satisfies vacuosuly formulas it does not cover, i.e., for any sentence

$$L \vDash S \Leftrightarrow (S \notin \boldsymbol{E}^-[L] \vee S \in L). \tag{2.10}$$

A *semikernel model* of a theory Γ is a (language graph with a) semikernel containing Γ. For the use with the sequent calculus, the following formulation adapts this definition. A sequent $\Gamma \Rightarrow \Delta$ is valid, $\Gamma \vDash \Delta$, iff for every language graph $G = G_M(\mathcal{L})$, every semikernel L of G satisfies it, $L \vDash \Gamma \Rightarrow \Delta$, under every valuation $\alpha \in (M^{o\mathcal{V}} \times (\boldsymbol{S}_M)^{s\mathcal{V}})$ of free variables of Γ, Δ, $L \vDash_\alpha \Gamma \Rightarrow \Delta$, i.e.:

$$\alpha(\Gamma) \cup \alpha(\Delta) \subseteq \boldsymbol{E}^-[L] \Rightarrow \alpha(\Gamma) \cap \boldsymbol{E}^-(L) \neq \varnothing \vee \alpha(\Delta) \cap L \neq \varnothing \tag{2.11}$$

By the condition in the antecedent, L satisfies every sequent it does not cover. A kernel model is a special case when semikernel L happens to be a kernel. The covering condition is then trivially satisfied. Semikernels provide a model for reasoning that explodes from contradictions but not from paradoxes. A contradiction has namely no semikernel model. For instance, no semikernel of the graph $\wedge \to \neg A \to A...$ contains its source vertex $A \wedge \neg A$. The framed sentences in Example 2.5, on the other hand, form a semikernel model of the liar K. It is only partial and cannot be extended to any kernel, that is, to a consistent interpretation of the whole language respecting the liar's claim, [11, 12].

2.1. Reasoning in $\mathcal{L}^{\doteq}$. This paper does not focus on paradoxes but on the definition of truth with the intended interpretation of s-equality. This interpretation is obtained by excluding (semi)kernels containing $S \not\doteq S$ or $S \doteq T$ for distinct sentences S, T.

DEFINITION 2.12. *In language graphs $G_M(\mathcal{L}^{\doteq})$, the considered (semi)kernels contain neither negation of any trivial equality, $S \not\doteq S$, nor any equality $S \doteq T$ for syntactically distinct $S, T \in \boldsymbol{S}_M$.*

Speaking about all semikernels of a language graph for $\mathcal{L}^{\doteq}$, we mean only the ones satisfying this definition. Alternatively, we could remove edges $\boldsymbol{E}(S \doteq S, S \not\doteq S)$ and $\boldsymbol{E}(S \not\doteq T, S \doteq T)$ from the atomic 2-cycles, turning their sources into sinks of the graph (that belong to every covering semikernel). Theorem 2.6 does not apply under any such restrictions and the rest of this section establishes its counterpart, Theorem 2.20, for this semantics of $\mathsf{LSO}^{\doteq}$. The proof uses also the reasoning system $\mathsf{LSO}^{\doteq}$ that extends LSO with the following axiom schema and rules:

(uniAx) $\Gamma, A \doteq B \vdash \Delta$ – for any A, B not unifiable with occurs-check

$$\text{(rep)}\ \frac{A \doteq B, P[\phi\backslash B], \Gamma \vdash \Delta}{A \doteq B, P[\phi\backslash A], \Gamma \vdash \Delta} \qquad \text{(ref)}\ \frac{A \doteq A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta}$$

$$\text{(unif)}\ \frac{A \doteq B, P[\phi\backslash A] \doteq P[\phi\backslash B], \Gamma \vdash \Delta}{P[\phi\backslash A] \doteq P[\phi\backslash B], \Gamma \vdash \Delta}$$

As in LSO, possibly open formulas A, B are only legally substituted for free s-variables. Rules (rep) and (ref) follow [4], while (unif) reflects the unification process discharging the common context and extracting possibly distinct subexpressions from both sides of $\doteq$. Unification in (uniAx) is usual first-order, with the operator names acting as function symbols (occurs-check prevents, e.g., the unification of ϕ with $P(\phi)$.) Possible extensions of s-equality to other pieces of syntax, like open formulas, are not considered here. Syntactically, s-equalities may involve open expressions, e.g., $P(x) \doteq P(y)$ is true when both x and y are substituted by the same expression, but only $P(x) \doteq P(x)$, or closed $A \doteq A$, are universally valid. Like usual equality, s-equality satisfies the congruence axiom $A \doteq B \to P(A) \doteq P(B)$, derivable by (rep) followed by (ref). (Unif) allows to derive also the reverse implication $P(A) \doteq P(B) \to A \doteq B$, holding only for s-equality.

The following example illustrates reasoning with s-equality and the non-explosive character of the semikernel semantics.

Example 2.13. *Recall $K!Kl$ from Example 2.5: $Kl = \forall\phi(K\phi \to \neg\phi)$ and $\forall\phi(K\phi \to \phi \doteq Kl)$. The unresolvable odd cycle is now reflected by the provability of K always lying and not always lying. ((Contr)action is admissible.)*

$$\text{(1)}\qquad \dfrac{\dfrac{\text{(Contr)}\,\dfrac{(\forall^{\phi}_{L})\,\dfrac{Kl, K(Kl) \vdash K(Kl) \qquad \dfrac{Kl, K(Kl) \vdash Kl}{Kl, K(Kl), \neg Kl \vdash}}{Kl, K(Kl), K(Kl) \to \neg Kl \vdash}}{Kl, K(Kl), \forall\phi(K\phi \to \neg\phi) \vdash}}{K(Kl), \forall\phi(K\phi \to \neg\phi) \vdash}}{\dfrac{K(Kl), Kl \vdash}{K(Kl) \vdash \neg Kl}}$$

$$\dfrac{\dfrac{(\forall^{\phi}_{L})\,\dfrac{\dfrac{\dfrac{\text{(rep)}\,\dfrac{\vdots\,(1)}{K(Kl), \psi \doteq Kl, K\psi, Kl \vdash}}{K(Kl), \psi \doteq Kl, K\psi, \psi \vdash}}{K(Kl), \psi \doteq Kl, K\psi \vdash \neg\psi} \qquad K(Kl), K\psi \vdash \neg\psi, K\psi}{K(Kl), K\psi \to \psi \doteq Kl, K\psi \vdash \neg\psi}\,\text{legal } (K\phi \to \phi \doteq Kl)[\phi\backslash\psi]}{(\forall^{\phi}_{R})\,\dfrac{K(Kl), \forall\phi(K\phi \to \phi \doteq Kl) \vdash K\psi \to \neg\psi}{K(Kl), \forall\phi(K\phi \to \phi \doteq Kl) \vdash \forall\phi(K\phi \to \neg\phi)}\,\text{fresh } \psi}}{K(Kl), \forall\phi(K\phi \to \phi \doteq Kl) \vdash Kl}$$

As in natural reasoning, the liar lies and does not lie, but not much else follows. Semikernel model from Example 2.5 can be extended to one satisfying all $\mathcal{L}^{-}$ tautologies and no $\mathcal{L}^{-}$ contradictions. None of the latter are derivable from $K!Kl$.

As in the kernel semantics, $K!Kl \vDash Kl \land \neg Kl$, but since no semikernel can contain $Kl \land \neg Kl$, this means now that no semikernel containing $K!Kl$ covers $Kl \land \neg Kl$. $K!Kl$ itself is not a contradiction. A contradiction Γ entails $\Gamma \vdash \varnothing$ and has no semikernel model, while $K!Kl$ has one, which arises also from the infinite branch ∞ of the failed derivation of $K!Kl \vdash \varnothing$ (marks the repetitions of the contracted $\forall\phi(K\phi \to \phi \doteq Kl)$):*

$$\dfrac{\stackrel{\text{(uniAx)}}{K(Kl), S_1 \doteq Kl, * \vdash \varnothing} \qquad \dfrac{\dfrac{\stackrel{\text{(uniAx)}}{K(Kl), S_2 \doteq Kl, * \vdash K(S_1)} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\vdots\,\infty}{K(Kl), * \vdash K(S_1), K(S_2)}}{Kl \doteq Kl, K(Kl), * \vdash K(S_1), K(S_2)} \qquad \stackrel{\text{(Ax)}}{K(Kl), * \vdash K(S_1), K(S_2), K(Kl)}}{K(Kl), K(Kl) \to Kl \doteq Kl, * \vdash K(S_1), K(S_2)}}{K(Kl), * \vdash K(S_1), K(S_2)}\,(\forall^{\phi}_{L})}{K(Kl), K(S_2) \to S_2 \doteq Kl, * \vdash K(S_1)}}{K(Kl), * \vdash K(S_1)}\,(\forall^{\phi}_{L})}{\dfrac{K(Kl), K(S_1) \to S_1 \doteq Kl, * \vdash \varnothing}{K(Kl), \forall\phi(K\phi \to \phi \doteq Kl), * \vdash \varnothing}\,(\forall^{\phi}_{L})}$$

Since $K!Kl \nvdash \varnothing$ while $Kl \land \neg Kl \vdash A$ for every A, rule (cut) $\dfrac{K!Kl \vdash Kl \land \neg Kl \quad Kl \land \neg Kl \vdash A}{K!Kl \vdash A}$ is not admissible. Adding (cut) does not trivialise the logic, but only turns its variant exploding from contradictions but not from paradoxes, into one where also paradoxes entail everything [11, 12].

2.2. Consistency of $\mathcal{L}^{\doteq}$. $\mathsf{LSO}^{\doteq}$ is sound and complete for the semikernel semantics (2.11). The proofs have some novel elements due to the novelty of the language and semantics but, generally, follow the standard route and are given in the appendix.

FACT 2.14 (5.1). *If* $\Gamma \vdash \Delta$ *in* $\mathsf{LSO}^{\doteq}$ *then* $\Gamma \vDash \Delta$*, according to (2.11).*

FACT 2.15 (5.3). *For a countable* $\mathcal{L}^{\doteq}$*, if* $\Gamma \nvdash \Delta$ *in* $\mathsf{LSO}^{\doteq}$ *then there is a language graph* $G(\mathcal{L}^{\doteq})$ *with a semikernel* L *such that (i)* $\Gamma \subseteq L$ *and (ii)* $\Delta \subseteq \boldsymbol{E}^{-}(L)$.

The last fact is used frequently, providing a covering semikernel for an uprovable sequent. Unlike in the usual case, these two facts do not establish consistency of $\mathsf{LSO}^{\doteq}$. Soundness admits namely both $\varnothing \vdash A$ and $\varnothing \vdash \neg A$, for a sentence A that is not covered by any semikernel (if such exists). If this happens then no language graph possesses a kernel. Establishing consistency of $\mathsf{LSO}^{\doteq}$, this section shows also that such kernels exist for languages containing $\doteq$ with the intended interpretation.

A theory (set of formulas) Γ is *p-consistent* if $\Gamma \nvdash A$ or $\Gamma \nvdash \neg A$, for each sentence A, and *p-complete* if $\Gamma \vdash A$ or $\Gamma \vdash \neg A$ for each sentence A.[4] P-consistency has finite character: a proof of any A from possibly infinite Γ is finite. It is easy to see that $\mathsf{LSO}^{\doteq}$ does not prove the empty sequent, $\varnothing \nvdash \varnothing$, but this does not suffice for its p-consistency (that is, p-consistency of the empty theory $\varnothing$). As noted above, even soundness for semikernel semantics does not suffice, but it underlies the following proof of p-consistency. First we register a fact used in the proof.

LEMMA 2.16. *If* $\varnothing \vdash A$ *and* $\varnothing \vdash B$*, while a language graph* G *has semikernels* L_A *containing* A *and* L_B *containing* B*, then* G *has a semikernel containing* A *and* B.

PROOF. Let SK be all semikernels of G. For $L \in SK$ let its theory be $Th(L) = \{T \in \boldsymbol{S}^{\doteq} \mid L \vdash T\}$, and $Th(SK) = \bigcap_{L \in SK} Th(L)$.

For $L_A, L_B \in SK$, both $\neg A \notin Th(L_A)$ and $\neg B \notin Th(L_B)$ by soundness of $\mathsf{LSO}^{\doteq}$, so neither $\neg A$ nor $\neg B$ is in $Th(SK)$, while both A, B are. Since $A \in Th(SK)$ and $Th(SK) \nvdash \neg B$ so, by Fact 2.15, there is a semikernel L that contains A with $\neg B \in \boldsymbol{E}^{-}(L)$, which means that also $B \in L$. $\square$

Let $\mathbf{C}$ range over all, possibly empty, sets of trivial equations $C \doteq C$. For any Γ, Δ and any such $\mathbf{C}$, if $\Gamma \vdash \Delta$ then $\mathbf{C}, \Gamma \vdash \Delta$ (by weakening) and vice versa by (rep).

LEMMA 2.17. *For every formula* A *and any* $\mathbf{C}_1, \mathbf{C}_2$*, if* $\mathbf{C}_1 \vdash A$*, then (a)* $A, \mathbf{C}_2 \nvdash \varnothing$ *and (b)* A *is covered by a semikernel.*

PROOF. (b) follows from (a) by Fact 2.15 that provides a semikernel containing A whenever $A, \mathbf{C}_2 \nvdash \varnothing$. We show that $\mathbf{C}_1 \nvdash A$ or $A, \mathbf{C}_2 \nvdash \varnothing$ by supposing the contrary and proceeding by induction on the length of a shortest proof $\mathbf{C}_1 \vdash A$ (over all $\mathbf{C}_1$) and, secondarily, of a shortest proof $A, \mathbf{C}_2 \vdash \varnothing$.

1. (Axioms) By Definition 2.1 of language graphs, equations and inequations have covering semikernels at the atomic 2-cycles. By Definition 2.12, each $S \doteq S$ forms a (relevant) semikernel, and so does each $S \not\doteq T$ for syntactically distinct S, T. These semikernels cover also the dual literals.

The proofs of length 1, yielding a required format, use (uniAx) and give $B \doteq C, \mathbf{C}_2 \vdash \varnothing$ for non-unifiable B, C. Every semikernel covering (any instance of) $B \doteq C$ contains then (the respective instance of) $B \not\doteq C$. By soundness, Fact 2.14, $\varnothing \nvdash B \doteq C$, hence also $\mathbf{C}_1 \nvdash B \doteq C$ by (ref).

2. For any equation A, if $\mathbf{C}_1 \vdash A$ then A is valid by soundness of $\mathsf{LSO}^{\doteq}$, which implies $A \nvdash \varnothing$. Then also $A, \mathbf{C}_2 \nvdash \varnothing$ by (ref). Thus, with the first premise $\mathbf{C}_1 \vdash B \doteq C$, this argument excludes derivation of the second ending, e.g., with (rep) $\dfrac{B \doteq C, P[B] \doteq P[C], \mathbf{C}_2' \vdash \varnothing}{B \doteq C, \underbrace{P[B] \doteq P[B], \mathbf{C}_2'}_{\mathbf{C}_2} \vdash \varnothing}$.

For other proofs longer than 1, we consider the last rule in the supposed proof of $\mathbf{C}_1 \vdash A$. The second sequent can always happen to follow in the last step by (ref) $\dfrac{C \doteq C, \mathbf{C}_2, A \vdash \varnothing}{\mathbf{C}_2, A \vdash \varnothing}$. Such cases follow trivially by IH on the second induction parameter and are not mentioned.

[4] The prefix "p-" abbreviates "provability". P-consistency entails the semantic consistency – the existence of kernels in language graphs for $\mathcal{L}^{\doteq}$. This, however, is the main Theorem 2.20 of this section and until it is shown distinguishing the two may be helpful. P-completeness is the usual completeness of a theory, only highlighting its difference from the completeness of a reasoning system.

3. If $\dfrac{A, \mathbf{C}_1 \vdash \varnothing}{\mathbf{C}_1 \vdash \neg A}$ and $\dfrac{\mathbf{C}_2 \vdash A}{\neg A, \mathbf{C}_2 \vdash \varnothing}$ then one of the two assumptions fails by IH, and some semikernel L covers A. If $A \in L$ then L covers $\neg A$, while if $A \in \boldsymbol{E}^{-}(L)$, then adding $\neg A$ to L yields a semikernel covering $\neg A$.

4. Suppose that both

$$(\wedge_R)\ \dfrac{\mathbf{C} \vdash A \quad \mathbf{C}' \vdash B}{\mathbf{C}_1 \vdash A \wedge B} \text{ and } \dfrac{A, B, \mathbf{C}_2 \vdash \varnothing}{A \wedge B, \mathbf{C}_2 \vdash \varnothing}\ (\wedge_L).$$

By IH, both $A \nvdash \varnothing$ and $B \nvdash \varnothing$. Fact 2.15 yields semikernels L_A containing A and L_B containing B. Lemma 2.16 yields then a semikernel containing both A and B. By soundness, Fact 2.14, every semikernel covering $A \wedge B$ contains it. If $A \wedge B \vdash \varnothing$ then, by soundness of $\mathsf{LSO}^{\doteq}$, no semikernel covering $A \wedge B$ contains it, contradicting the previous sentence.

5. Suppose that $(\forall^{\phi}_R)\dfrac{\mathbf{C}_1 \vdash F\psi}{\mathbf{C}_1 \vdash \forall\phi F\phi}$, with fresh $\psi \in \Theta$, and $(\forall^{\phi}_L)\dfrac{F(S), \mathbf{C}_2 \vdash \varnothing}{\forall\phi F\phi, \mathbf{C}_2 \vdash \varnothing}$.

The proof $\mathbf{C}_1 \vdash F\psi$ is shorter than the proof $\mathbf{C}_1 \vdash \forall\phi F\phi$. Since (Ax) does not require atoms, also (a copy of) the former with S substituting ψ, $\mathbf{C}_1 \vdash F(S)$, is shorter than the latter. If also $F(S), \mathbf{C}_2 \vdash \varnothing$ then this contradicts IH.

6. Applications of (unif) yielding $\mathbf{C}_1$ in the antecedent of the conclusion, have also some $\mathbf{C}'_1$ in the antecedent of the premise. $\dfrac{B \doteq C, P[C] \doteq P[B], \mathbf{C}_0, ... \vdash ...}{P[C] \doteq P[B], \mathbf{C}_0, ... \vdash ...}$ If $P[C]$ is syntactically identical to $P[B]$, i.e., $\mathbf{C}_1$ in the conclusion is $P[C] \doteq P[C], \mathbf{C}_0$, then also B and C are identical and $\mathbf{C}'_1$ in the premise is $C \doteq C, \mathbf{C}_1$. IH applies then to the premise, whether for $... \vdash A$ or $..., A \vdash \varnothing$. □

The relevant special case gives p-consistency of $\mathsf{LSO}^{\doteq}$: if $\varnothing \vdash A$ then $A \nvdash \varnothing$, hence also $\varnothing \nvdash \neg A$.

We now reach the main claim of this section: the semantic consistency of $\mathcal{L}^{\doteq}$, witnessed by the existence of kernels in its language graphs. For a finite set of sentences, if S denotes their conjunction, then either $\varnothing \nvdash S$ or $\varnothing \vdash S$, hence $\varnothing \nvdash \neg S$ by Lemma 2.17. In each case, Fact 2.15 gives a semikernel covering S. Every finite set of sentences is thus covered by a semikernel. Lemma 7.19 from [12] gives then a language graph $G(\mathcal{L}^{\doteq})$ with a kernel. Now, we can also ensure that every language graph for $\mathcal{L}^{\doteq}$ has a kernel. We show Lindenbaum's lemma for $\mathcal{L}^{\doteq}$, and that an p-complete p-consistent theory determines a kernel model.

Fact 2.18. *Every p-consistent theory Γ in $\mathcal{L}^{\doteq}$ has an p-complete p-consistent extension $\overline{\Gamma}$.*

Proof. We obtain $\overline{\Gamma}$ in the standard way: enumerate all sentences $S_1, S_2, ...$ of $\mathcal{L}^{\doteq}$, let $\Gamma_0 = \Gamma$ and

$$\Gamma_{i+1} = \begin{cases} \Gamma_i, S_i & \text{if } \Gamma_i \nvdash \neg S_i \\ \Gamma_i & \text{if } \Gamma_i \vdash \neg S_i \end{cases} \quad \text{and } \overline{\Gamma} = \textstyle\bigcup_{i \in \omega} \Gamma_i.$$

(If $\mathcal{L}_M$ is uncountable, we well-order the sentences and take unions in the limits.) By syntactic compactness $\overline{\Gamma}$ is p-consistent if every Γ_i is, and the latter follows since $\Gamma_i, S_i \vdash B$ and $\Gamma_i, S_i \vdash \neg B$ entail $\Gamma_i \vdash \neg S_i$. Namely, if $\Gamma_i, S_i \vdash X$ then every semikernel containing Γ_i, S_i contains X. As no semikernel contains both X and $\neg X$, if $\Gamma_i, S_i \vdash B$ and $\Gamma_i, S_i \vdash \neg B$, then no semikernel contains Γ_i, S_i. If $\Gamma_i, S_i \nvdash \varnothing$, then there is such a semikernel, hence $\Gamma_i, S_i \vdash \varnothing$, and then $\Gamma_i \vdash \neg S_i$. □

Given any language $\mathcal{L}^{\doteq}$ and any domain M, let $\mathcal{L}^{\doteq}_M$ contain a constant c_e for each $e \in M$, and Th be all provable sentences over $\mathcal{L}^{\doteq}_M$. Th is p-consistent by Lemma 2.17, so it has an p-complete p-consistent extension. This, in turn, determines a kernel model of $G_M(\mathcal{L}^{\doteq}_M)$, hence also of the isomorphic graph $G_M(\mathcal{L}^{\doteq})$.

Fact 2.19. *For a domain M and language $\mathcal{L}^{\doteq}_M$, with a name for each element of M, any p-complete p-consistent theory Γ over $\mathcal{L}^{\doteq}_M$ determines a kernel $K = \{S \mid \Gamma \vdash S\} \in Ker(G_M(\mathcal{L}^{\doteq}))$.*

Proof. Kernel condition (2.2) is ensured for each kind of vertex. By the invertibility of a rule we refer here to the provability of its premises when the conclusion is provable.

1. For $A \in \boldsymbol{A}_M$, exactly one of $A \in K$ or $\neg A \in K$, so $A \in K \Leftrightarrow \boldsymbol{E}(A) = \neg A \notin K$.
2. For each $\neg F$, p-consistency and p-completeness ensure that $\neg F \in K \Leftrightarrow \boldsymbol{E}(\neg F) = F \notin K$.

3. For a conjunction $(A\wedge B)\in K \Leftrightarrow \{A,B\}\subset K$ by rule $(\wedge_R)$ and its invertibility. P-consistency and p-completeness yield then the last equality: $(A\wedge B)\in K \Leftrightarrow \boldsymbol{E}(A\wedge B)=\{\neg A,\neg B\}\cap K=\varnothing$.
4. $\forall xFx\in K \Leftrightarrow F(t)\in K$ for each closed term t by (invertibility of) $(\forall_R)$. By consistency, $\boldsymbol{E}(\forall xFx)=\{\neg F(t)\mid t\in \boldsymbol{T}_M\}\cap K=\varnothing$. The same argument for $(\forall^{\phi}_R)$, which involves in addition $\forall\phi F\phi\in K\Leftrightarrow F\psi\in K$ for the fresh ψ used in the proof of $\forall\phi F\phi$. In the usual way, this proof gives a proof of $F(S)$ for every sentence S. □

In particular, every (FOL) language extended by s-quantifiers and $\doteq$ with the intended interpretation is consistent – it contains no paradoxes, no A and $\neg A$ is provable, while its graphs possess kernels. This theorem, missing in [11, 12], is fundamental for the following development.

Theorem 2.20. *Each language graph $G_M(\mathcal{L}^{\doteq})$ has a kernel with the intended interpretation of $\doteq$.*

§3. Truth definition. We now let $\mathcal{L}$ stand for a language containing s-equality $\doteq$ as the only operator, and an underlying (first-order) language $\mathcal{L}^-$. By $G=(\boldsymbol{S},\boldsymbol{E})$ we denote the language graph for $\mathcal{L}$ over any domain M. An appropriately axiomatized new binary operator E, in language $\mathcal{L}^{\mathsf{E}}=\mathcal{L}\cup\{\mathsf{E}\}$, represents the edge relation $\boldsymbol{E}^{\mathsf{E}}$ of the language graph $G^{\mathsf{E}}=(\boldsymbol{S}^{\mathsf{E}},\boldsymbol{E}^{\mathsf{E}})$ for $\mathcal{L}^{\mathsf{E}}$, which contains G as an induced subgraph. The restriction of edge relation $\boldsymbol{E}^{\mathsf{E}}$ to G is the latter's edge relation, $\boldsymbol{E}^{\mathsf{E}}\cap(\boldsymbol{S}\times\boldsymbol{S})=\boldsymbol{E}$. Truth is then defined by the kernel condition (2.3) in the language $\mathcal{L}^{\mathsf{T}}=\mathcal{L}^{\mathsf{E}}\cup\{\mathsf{T}\}$, with a new graph $G^{\mathsf{T}}=(\boldsymbol{S}^{\mathsf{T}},\boldsymbol{E}^{\mathsf{T}})$. We have thus three graphs $G\subset G^{\mathsf{E}}\subset G^{\mathsf{T}}$, over domain M, with the inclusions marking the induced-subgraph relation. The truth theory is interpreted in G^{T}, as the truth operator addresses all sentences of $\mathcal{L}^{\mathsf{T}}$, not only of $\mathcal{L}$.

The edge relation of G^{E} is captured by the operator E defined by the axioms below, reflecting the cases from Definition 2.1. (The unary arguments abbreviate arbitrary arities.)

Definition 3.1. *Theory Ed over $\mathcal{L}^{\mathsf{E}}$ consists of axiom schemas 1, 4 and axioms 2, 3:*

$$\begin{array}{lrlcll}
1a. & \forall x\forall\theta: & \mathsf{E}(Px,\theta) & \leftrightarrow & \theta\doteq\neg Px & \text{for each predicate } P\\
1b. & \forall\phi,\theta: & \mathsf{E}(O\phi,\theta) & \leftrightarrow & \theta\doteq\neg O\phi & \text{for each operator } O\\
2. & \forall\alpha,\theta: & \mathsf{E}(\neg\alpha,\theta) & \leftrightarrow & \theta\doteq\alpha & \\
3. & \forall\alpha,\beta,\theta: & \mathsf{E}(\alpha\wedge\beta,\theta) & \leftrightarrow & (\theta\doteq\neg\alpha)\vee(\theta\doteq\neg\beta) & \\
4a. & \forall z\forall\theta: & \mathsf{E}(\forall x.A(z,x),\theta) & \leftrightarrow & \exists y(\theta\doteq\neg A(z,y)) & \text{for each } \forall x.A(z,x)\in\boldsymbol{S}^{\mathsf{E}}_z\\
4b. & \forall z\forall\theta: & \mathsf{E}(\forall\phi.A(z,\phi),\theta) & \leftrightarrow & \exists\psi(\theta\doteq\neg A(z,\psi)) & \text{for each } \forall\phi.A(z,\phi)\in\boldsymbol{S}^{\mathsf{E}}_z
\end{array}\tag{3.1}$$

Axioms 4 admit formulas with free variables z, but they apply equally to sentences $\forall xAx$. A special case of 1b, for $O\phi=\mathsf{E}(\alpha,\beta)$, is axiom $\forall\alpha,\beta,\theta:\mathsf{E}(\mathsf{E}(\alpha,\beta),\theta)\leftrightarrow\theta\doteq\neg\mathsf{E}(\alpha,\beta)$. In kernels of any G^{E}_M containing Ed, E-vertices represent the edge relation $\boldsymbol{E}^{\mathsf{E}}$, in the following sense.

Fact 3.2. *For any M, if $Ed\subseteq K\in Ker(G^{\mathsf{E}}_M)$ then $\forall A,B\in\boldsymbol{S}^{\mathsf{E}}_M:\mathsf{E}(A,B)\in K\Leftrightarrow\boldsymbol{E}^{\mathsf{E}}(A,B)$.*

Proof. The same pattern is used for each pair of the corresponding points from Definitions 2.1 and 3.1. For any graph G^{E}_M, and $K\in Ker(G^{\mathsf{E}})$ with $Ed\subseteq K$:

1. For each $m\in M$ and atom Pm: $\boldsymbol{E}^{\mathsf{E}}(Pm,\theta)\overset{2.1.1}{\iff}\theta\doteq\neg Pm\overset{3.1.1}{\iff}\mathsf{E}(Pm,\theta)\in K$.
2. For each $\neg A\in\boldsymbol{S}_M:\boldsymbol{E}^{\mathsf{E}}(\neg A,\theta)\overset{2.1.2}{\iff}\theta\doteq A\overset{3.1.2}{\iff}\mathsf{E}(\neg A,\theta)\in K$.
3. For each $A,B\in\boldsymbol{S}_M:\boldsymbol{E}^{\mathsf{E}}(A\wedge B,\theta)\overset{2.1.2}{\iff}\theta\doteq\neg A$ or $\theta\doteq\neg B\overset{3.1.3}{\iff}\mathsf{E}(A\wedge B,\theta)\in K$.

4a. For each $n\in M:\boldsymbol{E}^{\mathsf{E}}(\forall xA(n,x),\theta)\overset{2.1.4a}{\iff}\theta\doteq\neg A(n,m)$, for some $m\in M$ $\overset{3.1.4b}{\iff}\mathsf{E}(\forall xA(n,x),\theta)\in K$.

4b. For each $n\in M:\boldsymbol{E}^{\mathsf{E}}(\forall\phi A(n,\phi),\theta)\overset{2.1.4b}{\iff}\theta\doteq\neg A(n,S)$, for some $S\in\boldsymbol{S}_M$ $\overset{3.1.4a}{\iff}\mathsf{E}(\forall xA(n,\phi),\theta)\in K$. □

This fact, together with (2.3), implies that for every $K\in Ker(G^{\mathsf{E}}_M)$ and sentence $S\in\boldsymbol{S}^{\mathsf{E}}_M$, if $(G^{\mathsf{E}}_M,K)\vDash Ed$ then $\big((G^{\mathsf{E}}_M,K)\vDash S$ iff $(G^{\mathsf{E}},K)\vDash\forall\theta(\mathsf{E}(S,\theta)\to\neg\theta)\big)$, yielding

$(G^{\mathsf{E}},K)\vDash\forall\phi(\phi\leftrightarrow\forall\theta(\mathsf{E}(\phi,\theta)\to\neg\theta))$.

We expand language $\mathcal{L}^{\mathsf{E}}$ to $\mathcal{L}^{\mathsf{T}}$ adding the unary operator T, and add the following axiom to Ed:

$$\forall\phi(\mathsf{T}\phi \leftrightarrow \forall\theta(\mathsf{E}(\phi,\theta) \to \neg\theta)). \tag{3.3}$$

Theory $Tr \supset Ed$ extends also axioms 4 to formulas in $\boldsymbol{S}_z^{\mathsf{T}}$ and axiom 1b to operator T, i.e.,

$$\forall\phi,\theta(\mathsf{E}(\mathsf{T}\phi,\theta) \leftrightarrow \theta \doteq \neg\mathsf{T}\phi). \tag{3.4}$$

The consistency of Tr is not immediate, but it follows due Ed and Tr introducing E and T essentially by definitional extensions, as shown below in Theorem 3.8. Convention T holds now for all, also open, formulas.

Fact 3.5 (Convention T). *For each $\mathcal{L}^{\mathsf{T}}$ formula $F(x,\phi)$ with free variables x,ϕ and each $K \in Ker(G_M^{\mathsf{T}})$ with $Tr \subseteq K$: $(G_M^{\mathsf{T}},K) \vDash \forall x,\phi(F(x,\phi) \leftrightarrow \mathsf{T}(F(x,\phi)))$.*

Proof. (a) For any sentence $A \in \boldsymbol{S}_M^{\mathsf{T}}$, the equivalence $\mathsf{T}(A) \leftrightarrow \forall\theta(\mathsf{E}(A,\theta) \to \neg\theta)$ follows by (3.3). (b) Now, $\mathsf{E}(A,B) \Leftrightarrow \boldsymbol{E}^{\mathsf{T}}(A,B)$ holds for each pair $A,B \in \boldsymbol{S}_M^{\mathsf{E}}$ in G_M^{E} by Fact 3.2, and in G_M^{T} since G_M^{E} is its induced subgraph. For atoms $\mathsf{E}(\mathsf{T}\phi,\theta)$ this equivalence holds by (3.4), and for the composite sentences with such atoms by the argument from Fact 3.2.

For each sentence $F(m,S) \in \boldsymbol{S}_M^{\mathsf{T}}$, this gives the first two equivalences below. The remaining ones follow by (2.3) and the properties of kernels and language graphs.

$$\begin{aligned}
(G^{\mathsf{T}},K) \vDash \mathsf{T}(F(m,S)) &\overset{(a)}{\Leftrightarrow} (G^{\mathsf{T}},K) \vDash \forall\theta(\mathsf{E}(F(m,S),\theta) \to \neg\theta)\\
&\overset{(b)}{\Leftrightarrow} \forall\theta(\boldsymbol{E}^{\mathsf{T}}(F(m,S),\theta) \to \neg\theta \in K)\\
&\Leftrightarrow \forall\theta(\boldsymbol{E}^{\mathsf{T}}(F(m,S),\theta) \to \theta \notin K)\\
&\Leftrightarrow F(m,S) \in K \Leftrightarrow (G^{\mathsf{T}},K) \vDash F(m,S)
\end{aligned}$$

Abstracting away the arguments of F yields $(G^{\mathsf{T}},K) \vDash \forall x,\phi(F(x,\phi) \leftrightarrow \mathsf{T}(F(x,\phi)))$. □

Theory Tr, implying in every $G \subset G^{\mathsf{T}} \in \mathcal{LG}r(\mathcal{L}^{\mathsf{T}})$ unrestricted Convention T also for the open formulas, should not be confused with the AST theory containing instances of [T] as axioms, called UTB. Of course, in AST some instances must be excluded (typically, some invoving truth predicate). In the operator setting all can be satisfied by the (basically identity) operator defined by (T). Unlike such a trivial operator, definition (3.3), involving 3.1, is quite substantial, capturing both the semantic graph structure and the intended definition of truth (2.3). Still, by Fact 3.5, operator T is fully transparent, even redundant for describing any specific issue. It makes each sentence S intersubstitutable with $\mathsf{T}(S)$. As in natural discourse, emphasising that one is telling the truth, $\mathsf{T}(S)$, does not add anything to saying S.

Theory Tr implies also equivalences of Tarski's truth definition, according to which T commutes with the logical connectives and quantifiers. In AST, any theory with these axioms (called "compositional truth", CT) must restrict them, typically, to the underlying language of arithmetic. In LSO truth is type-free and the truth operator distributes over all connectives and quantifiers.

Fact 3.6 (Compositionality). *For each $K \in Ker(G)$ with $Tr \subseteq K$, T distributes over logical connectives and quantifiers in (G,K).*

Proof. All cases follow directly from Fact 3.5. They hold for all formulas $A, B, Ax/A\phi$.

1. $(G,K) \vDash \forall x,\phi(\mathsf{T}(At(x,\phi)) \overset{3.5}{\leftrightarrow} At(x,\phi))$ – for all atoms $At(x,\phi)$
2. $(G,K) \vDash \mathsf{T}(\neg A) \overset{3.5}{\leftrightarrow} \neg A \leftrightarrow \neg\mathsf{T}(A)$ (the latter since $A \overset{3.5}{\leftrightarrow} \mathsf{T}(A)$)
3. $(G,K) \vDash \mathsf{T}(A \wedge B) \overset{3.5}{\leftrightarrow} A \wedge B \overset{3.5}{\leftrightarrow} \mathsf{T}(A) \wedge \mathsf{T}(B)$
4. The cases for both kinds of quantifers are virtually identical. The two tautologies:
   $\vDash \forall x(Ax \leftrightarrow \mathsf{T}(Ax)) \to (\forall x Ax \to \forall x\mathsf{T}(Ax))$ and
   $\vDash \forall x(Ax \leftrightarrow \mathsf{T}(Ax)) \to (\forall x\mathsf{T}(Ax) \to \forall x Ax)$, yield
   $(G,K) \vDash \forall x\mathsf{T}(Ax) \leftrightarrow \mathsf{T}(\forall x Ax)$, since $(G,K) \vDash \forall x(Ax \leftrightarrow \mathsf{T}(Ax))$ by Fact 3.5. □

These equivalences, expressing the truth value of a composite sentence in terms of the truth values of its (instantiated) subformulas, reflect the fact that the value of each vertex in a graph is determined by the values of its out-neighbours which, transitively, give all (instantiated) subformulas of a composite sentence. This holds, however, only locally and does not exclude circular dependencies which permeat through the subgraph of the metalanguage. Unlike in the usual well-founded languages, a sentence like $\forall\phi.\phi$ has itself as an instance, so $\mathsf{T}(\forall\phi.\phi) \Leftrightarrow \forall\phi(\mathsf{T}\phi)$ leads to $\mathsf{T}(\forall\phi.\phi)$ reappearing as one of the instances of the right side. The well-defined value of $\forall\phi.\phi$ makes also $\mathsf{T}(\forall\phi.\phi)$ well-defined – $\mathbf{0} \overset{2.4}{\Leftrightarrow} \forall\phi.\phi \overset{3.5}{\Leftrightarrow} \mathsf{T}(\forall\phi.\phi)$. It would be hard, however, to claim that $\mathsf{T}(\forall\phi.\phi) \Leftrightarrow \forall\phi\mathsf{T}(\phi)$ reduces the value of the more complex left side to the simpler right side where, after all, we have to evaluate T over all sentences. Compositionality interacts thus with circularity and, facing the sentential quantifiers, does not amount to recursively simplifying the more complex to the simpler.

Internalizing the graph and adding the truth operator amounts to a conservative extension. Stating this, the following theorem establishes also consistency of Tr by showing that it forms essentially a definitional extension. The form of Ed axioms in Definition 3.1 suggests this but, being schematic in the first argument, does not conform fully to the required format (2.7). Also axiom (3.4) does not quite fit this format. Hence the qualification "essentially" and the need for some additional arguments.

The proof of the theorem uses the following lemma (assuming AC), which is of general interest allowing to establish existence of kernels for limit interpretations of operators, where an *interpretation of an operator* O is the set of sentences about which O is true, $\{S \in \boldsymbol{S}_M \mid O(S) = \mathbf{1}\}$.

Lemma 3.7. *Let O be an operator in a language $\mathcal{L}$, $G \in Gr_M(\mathcal{L})$, $\{O_i \subseteq \boldsymbol{S}_M \mid i \in I\}$ and $O_J = \bigcup_{i\in J} O_i$ for $J \subseteq I$. If for each finite $J \in \mathcal{P}^{fin}(I)$, G has a kernel interpreting O as O_J, then G has a kernel interpreting O as O_I.*

Proof. For $Y \subseteq \boldsymbol{S}_M$, denote $O(Y) = \{O(S) \in \boldsymbol{S}_M \mid S \in Y\}$, while for each $S \in \boldsymbol{S}_M$ and $i \in I$

$ker_S = \{K \in Ker(G) \mid S \in K\}$

$ker_i = \{K \in Ker(G) \mid O(O_i) \subseteq K\} = \bigcap_{S\in O(O_i)} ker_S$.

($O_i \subseteq O_j$ implies $ker_i \supseteq ker_j$.) By the assumption, for each finite $J \in \mathcal{P}^{fin}(I)$ there is a kernel containing $O(O_J)$. This implies that

- each $ker_i \neq \varnothing$, and
- ker_I has FIP, where $ker_I = \begin{cases} \{ker_i \mid i \in I\} \text{ if } O_I = \boldsymbol{S}_M \text{ and otherwise} \\ \{ker_i \mid i \in I\} \cup \{ker_{\neg O(S)} \mid S \in \boldsymbol{S}_M \setminus O_I\}. \end{cases}$

The last summand ensures the interpretation $= O_I$, and not only only $\supseteq O_I$, when $O_I \neq \boldsymbol{S}_M$.

Let U be an ultrafilter with $ker_I \subseteq U \subset \mathcal{P}(Ker(G))$, existing by the ultrafilter lemma, and let $R = \{S \in \boldsymbol{S}_M \mid ker_S \in U\}$. Following facts, holding for every $A, B \in \boldsymbol{S}_M$, are used below:

(a) $ker_A \in U \vee ker_{\neg A} \in U$ and $A \in R \vee \neg A \in R$, since $ker_{\neg A} = Ker(G) \setminus ker_A$.
(b) $ker_A \notin U \vee ker_{\neg A} \notin U$ and $A \notin R \vee \neg A \notin R$, since $ker_A \cap ker_{\neg A} = \varnothing$.
(c) $ker_{A\wedge B} \in U \Leftrightarrow ker_A \in U \wedge ker_B \in U$ since $ker_{A\wedge B} = ker_A \cap ker_B$.

1. For each $i \in I$ and $S \in O_i$, $ker_{\neg O(S)} \cap ker_i \subseteq ker_{\neg O(S)} \cap ker_{O(S)} \overset{(b)}{=} \varnothing$, hence $ker_i \in U$ implies $ker_{O(S)} \in U$, and thus $S \in O_I \Rightarrow O(S) \in R$. If $O_I \neq \boldsymbol{S}_M$ then $ker_{\neg O(S)} \in U$ for each $S \in \boldsymbol{S}_M \setminus O_I$, and thus $S \notin O_I \Rightarrow \neg O(S) \in R$. Consequently $O(S) \in R \Leftrightarrow S \in O_I$.

2. We show $R \in Ker(G)$, verifying first its independence: $\boldsymbol{E}(S) \subseteq \boldsymbol{S}_M \setminus R$, for each $S \in R$.
(i) For each atom: $S \in R \Rightarrow \boldsymbol{E}(S) = \{\neg S\} \subset \boldsymbol{S}_M \setminus R$ by (b).
(ii) For each negation: if $\neg S \in R$ then $\boldsymbol{E}(\neg S) = \{S\} \subset \boldsymbol{S}_M \setminus R$ by (b).
(iii) For each conjunction: $\boldsymbol{E}(S_1 \wedge S_2) = \{\neg S_1, \neg S_2\}$, hence $ker_{S_1\wedge S_2} \cap ker_{\neg S_i} \overset{(b)(c)}{=} \varnothing$, for $i \in \{1,2\}$. Thus either $S_1 \wedge S_2 \notin R$ or both $\neg S_i \notin R$, i.e., $S_1 \wedge S_2 \in R \Rightarrow \boldsymbol{E}(S_1 \wedge S_2) \subset \boldsymbol{S}_M \setminus R$.
(iv) For $\forall$-quantified sentence: $\boldsymbol{E}(\forall x Fx) = \{\neg Fm \mid m \in M\}$, so $ker_{\forall xFx} \cap ker_{\neg Fm} = \varnothing$ for each $m \in M$, hence if $ker_{\forall xFx} \in U$ then $ker_{\neg Fm} \notin U$ for all $m \in M$, so that $\forall xFx \in R \Rightarrow \boldsymbol{E}(\forall xFx) \subset \boldsymbol{S}_M \setminus R$. (The same argument with $m \in \boldsymbol{S}_M$, for sententially quantified $\forall\phi F\phi$.)

3. Showing now $\boldsymbol{S}_M \setminus R \subseteq \boldsymbol{E}^-(R)$, i.e., that each $S \notin R$ has an edge to R, we obtain $R \in Ker(G)$.
(i) For each atom: $S \notin R \overset{(a)}{\Rightarrow} \boldsymbol{E}(S) = \{\neg S\} \subset R \Rightarrow S \in \boldsymbol{E}^-(R)$.
(ii) For each negation: $\neg S \notin R \overset{(a)}{\Rightarrow} \boldsymbol{E}(\neg S) = \{S\} \subset R \Rightarrow \neg S \in \boldsymbol{E}^-(R)$.
(iii) For $S_1 \wedge S_2 \notin R$: $ker_{(S_1\wedge S_2)} \notin U \overset{(c)}{\Leftrightarrow} (ker_{S_1} \notin U \vee ker_{S_2} \notin U) \overset{(a)}{\Leftrightarrow} (ker_{\neg S_1} \in U \vee ker_{\neg S_2} \in U)$. Hence $\neg S_1 \in R$ or $\neg S_2 \in R$ and $(S_1 \wedge S_2) \in \boldsymbol{E}^-(R)$ since $\boldsymbol{E}(S_1 \wedge S_2) = \{\neg S_1, \neg S_2\}$.

(iv) For $\forall xFx \notin R$, i.e., $ker_{\forall xFx} \notin U$: for each $m \in M$, $ker_{\forall xFx} \cap ker_{\neg Fm} = \varnothing$, since $\boldsymbol{E}(\forall xFx) = \{\neg Fm \mid m \in M\}$. Hence $ker_{\forall xFx} \in U \Leftrightarrow \forall m \in M(ker_{\neg Fm} \notin U)$, and thus

$\forall xFx \notin R \Leftrightarrow ker_{\forall xFx} \notin U \Leftrightarrow \exists m \in M(ker_{\neg Fm} \in U) \Leftrightarrow \exists m \in M(\neg Fm \in R)$,

yielding $\forall xFx \in \boldsymbol{E}^{-}(\neg Fm) \subset \boldsymbol{E}^{-}(R)$, for m with $\neg Fm \in R$.

(The same argument with $m \in \boldsymbol{S}_M$, for sententially quantified $\forall \phi F\phi$.) □

For language $\mathcal{L}^{\doteq}$ with the intended interpretation of the only operator $\doteq$, consistency is given by Theorem 2.20. For languages with specific interpretations of other operators, it requires an independent verification. The theorem below ensures that adding truth theory Tr (with fresh E, T), to a consistent theory over any language containing $\mathcal{L}^{\doteq}$, preserves consistency.

Theorem 3.8. *Extension of any theory over $\mathcal{L}^{\doteq}$ with E and T axiomatized by Ed and Tr is conservative:* $\forall \mathcal{L} \supseteq \mathcal{L}^{\doteq}, G \in \mathcal{L}\mathcal{G}r(\mathcal{L}), K \in Ker(G)\ \exists K^{\mathsf{T}} \in Ker(G^{\mathsf{T}}) : K^{\mathsf{T}} \supseteq K \cup Tr$.

Proof. Graph G, over domain M and language $\mathcal{L}^{\doteq}$, is an induced subgraph of graph G^{T} over $\mathcal{L}^{\mathsf{T}}$, expanding $\mathcal{L}^{\doteq}$ with E and T. Adding only these two symbols, without any axioms, is a definitional extension of whatever we started with, admitting an extension of any $K \in Ker(G)$ to the extended graph G^{T} by Theorem 2.7. We have to ensure this also with the axioms Ed and Tr.

Considering first Ed, we define partial interpretations of E for atoms, propositional connectives, and for each quantified $\mathcal{L}^{\mathsf{T}}$-sentence, by axioms reflecting those of Ed from Definition 3.1:

1*a*.	$\forall\gamma,\theta : \mathsf{E}_P(\gamma,\theta)$	$\leftrightarrow$	$\exists x((\gamma \doteq Px) \wedge (\theta \doteq \neg Px))$	– for each predicate P
1*b*.	$\forall\gamma,\theta : \mathsf{E}_O(\gamma,\theta)$	$\leftrightarrow$	$\exists\phi((\gamma \doteq O\phi) \wedge (\theta \doteq \neg O\phi))$	– for each operator O
2.	$\forall\gamma,\theta : \mathsf{E}_{\neg}(\gamma,\theta)$	$\leftrightarrow$	$\exists\alpha((\gamma \doteq \neg\alpha) \wedge (\theta \doteq \alpha))$	
3.	$\forall\gamma,\theta : \mathsf{E}_{\wedge}(\gamma,\theta)$	$\leftrightarrow$	$\exists\alpha,\beta((\gamma \doteq \alpha\wedge\beta) \wedge ((\theta \doteq \neg\alpha) \vee (\theta \doteq \neg\beta)))$	
4*a*.	$\forall\gamma,\theta : \mathsf{E}_{\forall xAzx}(\gamma,\theta)$	$\leftrightarrow$	$\exists z((\gamma \doteq \forall xA(z,x)) \wedge \exists x(\theta \doteq \neg A(z,x)))$	– for each $\forall xA(z,x) \in \boldsymbol{S}_z^{\mathsf{T}}$
4*b*.	$\forall\gamma,\theta : \mathsf{E}_{\forall\phi Az\phi}(\gamma,\theta)$	$\leftrightarrow$	$\exists z(\gamma \doteq \forall\phi A(z,\phi)) \wedge \exists\phi(\theta \doteq \neg A(z,\phi))$	– for each $\forall\phi A(z,\phi) \in \boldsymbol{S}_z^{\mathsf{T}}$

In axiom 1b, operators O come from the extended language $\mathcal{L}^{\mathsf{T}}$ (including (3.4)), hence there is a special case for operator E, namely, $\forall\gamma,\theta : \mathsf{E}_{\mathsf{E}}(\gamma,\theta) \leftrightarrow \exists\alpha,\beta(\gamma \doteq \mathsf{E}(\alpha,\beta)) \wedge (\theta \doteq \neg\mathsf{E}(\alpha,\beta))$.

Each axiom above introduces operator E_u, for some index u, by a definitional, hence conservative, extension (of any $\mathcal{L}^{\doteq}$ theory). Viewing it is an interpretation of E, the following argument shows conservativeness of E interpreted as the union of all such approximations.

Let U be the set of all indices of these operators (i.e., all predicate and operator symbols, $\neg, \wedge$, and all quantified formulas $\boldsymbol{S}_z^{\mathsf{T}}$), and R_u, for $u \in U$, denote the right side of the axiom for E_u above. For any finite subset $C \in \mathcal{P}^{fin}(U)$, $\mathsf{E}(\gamma,\theta) \leftrightarrow \bigvee_{c\in C} R_c(\gamma,\theta)$ is a definitional extension (of any theory over $\mathcal{L}^{\doteq}$ which may be modelled by K). By Theorem 2.8, any $K \in Ker(G)$ can be then extended to a kernel K^C of G^{T}, with $\mathsf{E}(A,B) \in K^C$ iff $R_c(A,B) \in K$ for some $c \in C$, i.e., with the interpretation of E equal to the union of the interpretations of $\mathsf{E}_c, c \in C$. By Lemma 3.7, graph G^{T} has then a kernel K^{E}, containing K, with E interpreted by the union of all these interpretations, i.e., satisfying axioms Ed from Definition 3.1.

Now, the language $\mathcal{L}^{\mathsf{T}}$ of graph G^{T} contains operator T, but so far it remains unrestricted and enters only into the arguments of E. The conservative extension Ed introduces only operator E but does not restrict T in any way. Axiom (3.3) fits now the form of a proper definitional extension – of the theory Ed, formulated already in the language $\mathcal{L}^{\mathsf{T}}$. By Theorem 2.8, kernel K^{E} of G^{T} can be extended to a kernel modelling also (3.3). □

A couple of points may deserve empahsis. Given a domain M, for a language $\mathcal{L}^{\doteq}$ with the names for all elements of M, theory Tr axiomatizes the truth of $\mathcal{L}^{\doteq}$ sentences in M and of all $\mathcal{L}^{\mathsf{T}}$ sentences in $G_M(\mathcal{L}^{\mathsf{T}})$. This theory, however, is relative to the actual truths of the atoms, that is, to the actual structure over M. These atoms are left undetermined by Tr. Given any $\mathcal{L}^{-}$ structure (M,μ), i.e., a FOL domain M with a valuation μ of $\mathcal{L}^{-}$ atoms, the concept of truth in this structure, modelled by $G_M(\mathcal{L}^{\mathsf{T}})$, is captured by Tr augmented with the axioms reflecting the valuation μ.

Another undetermined issue is the cardinality of M. Naming all its elements ensures that we can address all of M, but does not exclude possible confusions, where some names denote the same element. Like the specific valuations of atoms, this is left undetermined by Tr and dependent on the capacity of the object-language $\mathcal{L}^{-}$ to address it. If $\mathcal{L}^{-}$ contains the usual equality $=$, it may enforce distinct interpretations of distinct names. This does not seem to pertain to the general notion of truth, but to its specific instances, relative to the object-language $\mathcal{L}^{-}$.

§4. Some philosophical remarks. One may object that we have not treated syntax, except for the s-equality, while the undefinability theorem arises only with a sufficiently strong theory of syntax. This is indeed the case, but for truth modelled by a predicate, so that metalanguage must be coded in the object-language. Since LSO is its own metalanguage, no additional means are needed for such a self-representation. Using the truth operator applicable to sentences, instead of a truth predicate applicable to their names, the truth theory in LSO is relieved of the burden to code its syntax. The paradigmatic way of such a coding is by arithmetisation, so this implies also severing the bonds that connect truth theories with arithmetic in AST. Modelling of truth in LSO does not benefit from any arithmetic contributions, but neither suffers from any, like the diagonalisation lemma. The gain seems worth the price, at least philosophically. This does not in any way exclude the possibility of the traditional interweaving of the language with its arithmetic coding, which can be studied within the presented framework without any changes.

By Theorem 3.8, any consistent theory, in particular, any consistent theory of syntax can be extended consistently with the truth operator. The problem is thus placed where it belongs, at the formulation of an adequate theory of syntax, especially, since the challenge here is the substitution operator for formulas (that is not reducible to the substitution of terms). It can be handled in a consistent manner, but is left for the future work over an extension of LSO with such an operator and quantification over open formulas.

On a more philosophical note, the obtained theory of truth is, in a sense, deflationary; it does not add anything to any subject matter. The extension with Tr is conservative, while the unrestricted Convention T makes each sentence $\mathcal{T}(S)$ equivalent to – hence intersubstitutable with – S. The apparent hollowness of the notion of truth is the consequence of the unlimited validity of Convention T. Taking it as a condition of the material adequacy, this form of emptiness seems unavoidable.

We have disregarded Convention T as the definition of truth exactly for this reason. Unlike it, the presented definition seems quite substantial, in spite of its transparency and conservativeness. It defines the semantic graph structure and the relation over it that serves as the truth condition. Trivializing and simplifying, this relation could be expressed as saying: a sentence is true iff each sentence it negates is false. (Semantically, even an atomic sentence negates its negation.) This reminds about a coherence definition and the element of coherence, or rather holism, is strongly present. For instance, each two sententially quantified sentences in the graph have paths both ways, forming a connected component, where every sentence depends in some way on every other, even on itself. In this sense, truth in the metalanguage is holistic.

This holism is however only half of the story. The subgraph for the object-language is well-founded, except for the 2-cycles of dual literals at its leaves. It provides only an equivalent representation of the classical semantics, where each valuation of atoms induces unique values to all sentences. Truth at this level was viewed by Tarski as a form of correspondence, and no reason suggests to change this view, especially, if we compare it to the truth in the metalanguage.

The two perspectives complement each other, just as do the two language levels. Truth has, like science in Quine's formulation, "its double dependence upon language and experience" [10]. The part of the language which depends on the language displays some phenomena not occurring in the other part. The truth value of one sentence may depend on the truth value of another which, in turn, depends on the first. The sheer possibility of evaluating coherently all sentences of the metalanguage becomes uncertain and can be disturbed by paradoxical claims. Nothing of the sort occurs in the object-language. Truth of its statements, concerned exclusively with the nonlinguistic facts, is founded in the correspondence to such facts. It seems highly satisfactory that these two, apparently contrary views, appear now as expressions of one formal concept.

And how do we know that this concept of truth is true, that it is correct? Satisfaction of Convention T, after all, is not proven within LSO. Formalising the metaconcepts like proof or consistency in LSO is certainly an intriguing challenge for further work. Taking this question here less formally, do we not need some external verification that the language models what is really out there? Does the adequacy of a truth definition not require any measure by some external standard? Apparently it does, but this view can be reversed. We start with a language $\mathcal{L}^{\doteq}$, in which we define the operators E and T the way we did. These are just operators like any others. Then we note that E can be understood semantically, as the edge relation in the language graphs, while T as the feature of some sentences, namely, those we accept as true. We did not capture anything external, but only realized that our formalisation in $\mathcal{L}^{\doteq}$ can be interpreted in the non-linguistic world. Deciding which one comes first seems futile, since what can be said *about* the latter relies on what can be *expressed in* the former, while what the former is saying depends on how the latter is understood. The possible differences between the world and our understanding of it can be safely left for some deeper investigations. By our definition, truth in the metalanguage is essentially linguistic, but at the object-level only captures the interdependency of the language and the world, granting the irrefutable primacy to the atomic facts.

Finally, the ability to formalise in the language its semantics and truth does not in any way trivialise the issue of deciding the truth of specific sentences. Defining what truth is does not settle the question what is true, which remains as undecidable and problematic as it was before. This may be disappointing for the reductionists, hoping for a model replacing truth with some simpler concept that would be easier to ascertain. But the truth at the object-level remains incomplete without its holistic counterpart at the metalevel, which clarifies primarily the peculiarities of its self-referential ascriptions. Together, they provide an internal representation of the truth concept in the language and a general picture of which (sentences) may be involved in the verification of the truth of specific sentences. Which of them are actually true depends on the way the world is, but also on the claims we make, especially about our own claims.

§5. Appendix. One verifies easily that the rules of weakening and contraction are admissible in $\mathsf{LSO}^{\doteq}$ (without changing the heights of the derivations). Also, the system obtained from $\mathsf{LSO}^{\doteq}$ by replacing (Ax) and rules $(\forall^{\phi}_L)$, $(\forall_L)$ and (rep) by the versions listed below (retaining the principal formula of the conclusion in the premise) is equivalent, proving exactly the same sequents. Below, $\mathsf{LSO}^{\doteq}$ refers to the system with these primed versions of (Ax)' and the rules.

(Ax)' $\Gamma \vdash \Delta$ – Γ and Δ share some *atom* $\qquad (\forall_L)'\ \dfrac{F(t), \forall x F(x), \Gamma \vdash \Delta}{\forall x F(x), \Gamma \vdash \Delta}$ – legal $F[x\backslash t]$

(rep)' $\dfrac{P[\phi\backslash A], A \doteq B, P[\phi\backslash B], \Gamma \vdash \Delta}{A \doteq B, P[\phi\backslash B], \Gamma \vdash \Delta}$ $\qquad (\forall^{\phi}_L)'\ \dfrac{F(S), \forall \phi F(\phi), \Gamma \vdash \Delta}{\forall \phi F(\phi), \Gamma \vdash \Delta}$ – legal $F[\phi\backslash S]$

Both following proofs essentially repeat the respective proofs from [11], with minor corrections and adaptations to account for the slight differences between the current paper and [11].

Fact 5.1 (2.14). *If $\Gamma \vdash \Delta$ in $\mathsf{LSO}^{\doteq}$ then $\Gamma \vDash \Delta$, according to (2.11), and the rules are semantically invertible.*

Proof. Given an arbitrary language graph $G_M(\mathcal{L}^{\doteq})$, we show the claim by induction on the length of the proof $\Gamma \vdash \Delta$. Semantic invertibility follows by showing that every semikernel L covering (each) premise of a rule satisfies it, assuming this for the rule's conclusion.

1. (Ax) is valid for any valuation α, if a semikernel L covers $\alpha(\Gamma) \cup \alpha(\Delta)$ and contains $\alpha(\Gamma)$, then it obviously contains also $\alpha(\Gamma \cap \Delta)$. (uniAx) is valid for the intended interepretation of $\doteq$, since non unifiability of $A \doteq B \in \Gamma$ means exacty that no substitution yields identical $\alpha(A)$ and $\alpha(B)$, hence no semikernel satisfying Definition 2.12 contains $\alpha(A) \doteq \alpha(B)$.

2. $(\wedge_R)$. For soundness, assume $\Gamma \vDash \Delta, A_1$ and $\Gamma \vDash \Delta, A_2$, and let semikernel L cover the rule's conclusion, under a valuation α in a given domain. Assume that $\alpha(\Gamma) \cup \alpha(\Delta) \subseteq \boldsymbol{E}^{-}(L)$ and $\alpha(A_1 \wedge A_2) \in \boldsymbol{E}^{-}[L]$ – otherwise the conclusion is satisfied under α. It follows also if $\alpha(A_1 \wedge A_2) \in L$, so suppose $\alpha(A_1 \wedge A_2) \in \boldsymbol{E}^{-}(L)$. Since $\boldsymbol{E}(\alpha(A_1 \wedge A_2)) = \{\neg\alpha(A_1), \neg\alpha(A_2)\}$ so, for some $i \in \{1, 2\}$, $\neg\alpha(A_i) \in L$, and then $\alpha(A_i) \in \boldsymbol{E}^{-}(L)$, contradicting the assumption $\Gamma \vDash \Delta, A_i$.

For invertibility, let $\Gamma \vDash \Delta, A_1 \wedge A_2$ and L cover $\alpha(\Gamma \cup \Delta)$ and $\alpha(A_1)$ (or $\alpha(A_2)$). If (*) $\alpha(\Gamma) \subseteq L$ and $\alpha(\Delta \cup \{A_1\}) \subseteq \boldsymbol{E}^{-}(L)$, then $L' = L \cup \{\neg\alpha(A_1)\}$ is a semikernel, since $\boldsymbol{E}(\neg\alpha(A_1)) = \{\alpha(A_1)\} \subseteq \boldsymbol{E}^{-}(L) \subseteq \boldsymbol{E}^{-}(L')$. L' covers also $\alpha(A_1 \wedge A_2) \in \boldsymbol{E}^{-}(\neg\alpha(A_1))$. Thus L' covers the conclusion, while $\alpha(\Gamma) \cap \boldsymbol{E}^{-}(L') = \varnothing$ and $\alpha(\Delta \cup \{A_1 \wedge A_2\}) \cap L' = \varnothing$, so $L' \not\models \Gamma \Rightarrow \Delta, A_1 \wedge A_2$, contrary to $\Gamma \models \Delta, A_1 \wedge A_2$. Hence (*) fails, so $\alpha(\Gamma) \cap \boldsymbol{E}^{-}(L) \neq \varnothing$ or $\alpha(\Delta \cup \{A_1\}) \cap L \neq \varnothing$. This, together with the same argument for A_2, shows invertibility.

Valuations α of the free variables do not affect the argument, so covering by L below is to be taken relatively to a given α, which we do not mention, except for $(\forall_R)$ and $(\forall_R^{\phi})$.

3. $(\wedge_L)$. For soundness, assume $\Gamma, A_1, A_2 \models \Delta$, let semikernel L cover the rule's conclusion, $\Gamma \subseteq L$ and $\Delta \subseteq \boldsymbol{E}^{-}(L)$. If $A_1 \wedge A_2 \in L$, then $\boldsymbol{E}(A_1 \wedge A_2) = \{\neg A_1, \neg A_2\} \subseteq \boldsymbol{E}^{-}(L)$, so $\boldsymbol{E}(\{\neg A_1, \neg A_2\}) = \{A_1, A_2\} \subseteq L$, contradicting $\Gamma, A_1, A_2 \models \Delta$. Thus $A_1 \wedge A_2 \in \boldsymbol{E}^{-}(L)$ and $L \models \Gamma, A_1 \wedge A_2 \Rightarrow \Delta$.

For invertibility, assume $\Gamma, A_1 \wedge A_2 \models \Delta$, let semikernel L cover the rule's premise, and assume $\Gamma \subseteq L$ and $\Delta \subseteq \boldsymbol{E}^{-}(L)$. If $A_1, A_2 \in L$, which is the only way L can contradict $\Gamma, A_1, A_2 \models \Delta$, then $\{\neg A_1, \neg A_2\} \subseteq \boldsymbol{E}^{-}(L)$, and $L' = L \cup \{A_1 \wedge A_2\}$ is also a semikernel:

$$\boldsymbol{E}(L') = \boldsymbol{E}(L \cup \{A_1 \wedge A_2\}) = \boldsymbol{E}(L) \cup \boldsymbol{E}(\{A_1 \wedge A_2\}) \subseteq \boldsymbol{E}^{-}(L) \cup \{\neg A_1, \neg A_2\} \subseteq \boldsymbol{E}^{-}(L) \subseteq \boldsymbol{V} \setminus (L \cup \{A_1 \wedge A_2\}).$$

The last inclusion follows because $\boldsymbol{E}^{-}(L) \subseteq \boldsymbol{V} \setminus L$ and $A_1 \wedge A_2 \notin \boldsymbol{E}^{-}[L]$, since $A_1 \wedge A_2 \in L$ contradicts $\Gamma, A_1 \wedge A_2 \models \Delta$ (as $\Gamma \subseteq L$ and $\Delta \subseteq \boldsymbol{E}^{-}(L)$), while $A_1 \wedge A_2 \in \boldsymbol{E}^{-}(L)$ contradicts independence of L, implying $\neg A_i \in L$ (for $i = 1$ or $i = 2$), while $\neg A_i \in \boldsymbol{E}^{-}(L)$ since $A_i \in L$.

Since $L' \not\models \Gamma, A_1 \wedge A_2 \Rightarrow \Delta$ contradicts the assumption, either $A_1 \notin L$ or $A_2 \notin L$, and $L \models \Gamma, A_1 \wedge A_2 \Rightarrow \Delta$ as desired.

4. $(\neg_R)$. For soundness, assume $\Gamma, A \models \Delta$, let semikernel L cover the rule's conclusion, and assume $\Gamma \subseteq L$ and $\Delta \subseteq \boldsymbol{E}^{-}(L)$. If $\neg A \in L$, we are done, while if $\neg A \in \boldsymbol{E}^{-}(L)$ then $A \in L$, which contradicts the assumption, since now $\Gamma \cup \{A\} \subseteq L$ and $\Delta \subseteq \boldsymbol{E}^{-}(L)$.

For invertibility, assuming $\Gamma \models \Delta, \neg A$, let L cover the rule's premise, $\Gamma \subseteq L$ and $\Delta \subseteq \boldsymbol{E}^{-}(L)$. If $A \in L$ then $\neg A \in \boldsymbol{E}^{-}(L)$ and $L \not\models \Gamma \Rightarrow \Delta, \neg A$, contradicting the assumption. Hence $A \in \boldsymbol{E}^{-}(L)$, as required for $L \models \Gamma, A \Rightarrow \Delta$.

5. $(\neg_L)$. For soundness, assume $\Gamma \models \Delta, A$, let L cover the rule's conclusion, $\Gamma \subseteq L$ and $\Delta \subseteq \boldsymbol{E}^{-}(L)$. If $\neg A \in \boldsymbol{E}^{-}(L)$, we are done, while if $\neg A \in L$ then $A \in \boldsymbol{E}(\neg A) \subseteq \boldsymbol{E}^{-}(L)$, contradicting the assumption, since now $\Gamma \cup \{A\} \subseteq L$ and $(\Delta \cup \{A\}) \subseteq \boldsymbol{E}^{-}(L)$.

For invertibility, assume $\Gamma, \neg A \models \Delta$, let L cover the rule's premise, $\Gamma \subseteq L$ and $\Delta \subseteq \boldsymbol{E}^{-}(L)$. If $A \in \boldsymbol{E}^{-}(L)$ then $L' = L \cup \{\neg A\}$ is a semikernel, because L is and $\boldsymbol{E}(\neg A) = \{A\} \subseteq \boldsymbol{E}^{-}(L)$. But L' contradicts the assumption, so $A \in L$, as required for $L \models \Gamma \Rightarrow \Delta, A$.

6. $(\forall_L)$'. For soundness, assume $F(t), \Gamma, \forall x F(x) \models \Delta$ and let L cover the rule's conclusion. If $\forall x F(x) \notin L$, i.e., $\forall x F(x) \in \boldsymbol{E}^{-}(L)$, then $L \models \Gamma, \forall x F(x) \Rightarrow \Delta$. If $\forall x F(x) \in L$ then also $F(t) \in L$, since $\neg F(t) \in \boldsymbol{E}(\forall x F(x)) \subseteq \boldsymbol{E}^{-}(L)$ and $\boldsymbol{E}(\neg F(t)) = \{F(t)\}$. As L covers the premise, either $\Gamma \cap \boldsymbol{E}^{-}(L) \neq \varnothing$, since $F(t) \notin \boldsymbol{E}^{-}(L)$, or $\Delta \cap L \neq \varnothing$. Either case yields the claim for L, which was arbitrary, so $\Gamma, \forall x F(x) \models \Delta$.

For invertibility, if $\Gamma, \forall x F(x) \models \Delta$ and L covers the rule's premise, it covers also this conclusion. Satisfying it, L trivially satisfies the premise.

7. $(\forall_R)$. For soundness, let (*) $\Gamma \models \Delta, F(y)$ and L cover the rule's conclusion under a given assignment α to $\mathcal{V}(\Gamma, \Delta, \forall x F(x)) \not\ni y$. Assume also $\alpha(\Gamma) \subseteq L$ and $\alpha(\Delta) \subseteq \boldsymbol{E}^{-}(L)$. If $\alpha(\forall x F(x)) \notin L$ then $\alpha(\forall x F(x)) \in \boldsymbol{E}^{-}(L)$ and some $\alpha(\neg F(m)) \in L$, since $\boldsymbol{E}(\alpha(\forall x F(x))) = \{\alpha(\neg F(m)) \mid m \in M\}$. Extending α with $\alpha(y) = m$, we obtain $L \not\models_{\alpha} \Gamma \Rightarrow \Delta, F(y)$, contrary to (*). Thus, $\alpha(\forall x F(x)) \in L$ and $L \models_{\alpha} \Gamma \Rightarrow \Delta, \forall x F(x)$.

For invertibility, if $L \not\models_{\alpha} \Gamma \Rightarrow \Delta, F(y)$, for $\alpha(y) = m$, i.e., $\alpha(\Gamma) \subseteq L$, $\alpha(\Delta) \subseteq \boldsymbol{E}^{-}(L)$ and $\alpha(F(m)) \in \boldsymbol{E}^{-}(L)$, then $L' = L \cup \{\alpha(\neg F(m))\}$ is a semikernel, because L is and $\boldsymbol{E}(\alpha(\neg F(m))) = \{\alpha(F(m))\} \subseteq \boldsymbol{E}^{-}(L) \subseteq \boldsymbol{E}^{-}(L')$. L' covers the conclusion since $\alpha(\forall x F(x)) \in \boldsymbol{E}^{-}(\alpha(\neg F(m)))$, but $L' \not\models_{\alpha} \Gamma \Rightarrow \Delta, \forall x F(x)$.

8. $(\forall_L^{\phi})$'. The argument repeats that for $(\forall_L)$. For soundness, let $\Gamma, F(S), \forall \phi F(\phi) \models \Delta$ and L cover the rule's conclusion (under a fixed α). If $\forall \phi F(\phi) \notin L$ then $\forall \phi F(\phi) \in \boldsymbol{E}^{-}(L)$, yielding $L \models \Gamma, \forall \phi F(\phi) \Rightarrow \Delta$. If $\forall \phi F(\phi) \in L$ then also $F(S') \in L$, for each sentence $S' \in \boldsymbol{S}_M$, since $\neg F(S') \in \boldsymbol{E}(\forall \phi F(\phi)) \subseteq \boldsymbol{E}^{-}(L)$ and $\boldsymbol{E}(\neg F(S')) = \{F(S')\}$. Thus L covers also the premise, hence, either $\Gamma \cap \boldsymbol{E}^-(L) \neq \varnothing$, since $F(S) \notin \boldsymbol{E}^-(L)$, or $\Delta \cap L \neq \varnothing$. Either case yields the claim for L, which was arbitrary (as was α), so $\Gamma, \forall\phi F(\phi) \vDash \Delta$.

For invertibility, assume $\Gamma, \forall\phi F(\phi) \vDash \Delta$. If L covers the rule's premise, then it covers also this conclusion. Satisfying it, L trivially satisfies the premise.

9. $(\forall_R^\phi)$. For soundness, if $\Gamma \vdash \Delta, F(\psi)$ for a fresh $\psi \in \Theta$, then (a copy of) this proof yields also $\Gamma \vdash \Delta, F(S)$ for every sentence $S \in \boldsymbol{S}_M$. By IH, $L \cap (\{F(S)\} \cup \Delta) \neq \varnothing$ for every semikernel L containing Γ and covering $\Delta, F(S)$.

Let semikernel L cover the rule's conclusion. If $\forall\phi F(\phi) \in L$ then L satisfies it. If $\forall\phi F(\phi) \notin L$ then $\forall\phi F(\phi) \in \boldsymbol{E}^-(L)$, so some $\neg F(S) \in L$, since $\boldsymbol{E}(\forall\phi F(\phi)) = \{\neg F(S) \mid S \in \boldsymbol{S}_M\}$. Now L covers also $\Gamma \Rightarrow \Delta, F(S)$ and $F(S) \notin L$. Since $\Gamma \vDash \Delta, F(S)$, either $\Gamma \cap \boldsymbol{E}^-(L) \neq \varnothing$ or $\Delta \cap L \neq \varnothing$. In each case L satisfies the conclusion.

For invertibility, assume $\Gamma \vDash \Delta, \forall\phi F(\phi)$, and let L cover $\Gamma \Rightarrow \Delta, F(\psi)$ under $\alpha(\psi) = S \in \boldsymbol{S}_M$. Assume $\alpha(\Gamma) \subseteq L$ and $\alpha(\Delta) \subseteq \boldsymbol{E}^-(L)$, since otherwise $L \vDash_\alpha \Gamma \Rightarrow \Delta, F(\psi)$. Then $\forall\phi F(\phi) \in L$ and, as L is a semikernel, $\neg F(S) \in \boldsymbol{E}^-(L)$, so $F(S) \in L$ and $L \vDash_\alpha \Gamma \Rightarrow \Delta, F(\psi)$. Since α was arbitrary, $L \vDash \Gamma \Rightarrow \Delta, F(\psi)$.

10. S-equality rules are sound and invertible. For each formula A, each instance $A' \doteq A'$ in the premise of (ref) is satisfied in each (considered) semikernel (Definition 2.12), hence satisfaction of the premise and of the conclusion are equivalent. Rule (rep)', from [4], is sound and invertible for every congruence on formulas, while (unif) is such for the syntactc equality, for which $A \doteq B$ is equivalent to $P(A) \doteq P(B)$, for every context $P(...)$. □

The following simple consequence of Definition 2.1 is used in the proof of Fact 2.15 below.

Fact 5.2. *In any graph $G_M \in \mathcal{LG}r(\mathcal{L})$, the following relations hold between the form of a nonatomic sentence $X \in \boldsymbol{S}^+_M$ and forms of its out- and in-neighbours:*

1. $\boldsymbol{E}^-(X) = \{\neg X\}$ *– when X does not start with $\neg$,*
2. $\boldsymbol{E}^-(\neg X) = \{\neg\neg X\} \cup \{X \wedge S \mid S \in \boldsymbol{S}^+_M\} \cup ...$
   $\{\forall\phi.D(\phi) \mid \exists S \in \boldsymbol{S}^+_M : D(S) = X\} \cup \{\forall x.D(x) \mid \exists t \in \boldsymbol{T}_M : D(t) = X\}$
3. *when X does not start with $\neg$, then each out-neighbour of X does,*
4. $\boldsymbol{E}(\neg X) = \{X\}$.

For atomic X, $\boldsymbol{E}^-(X) = \{\neg X\} = \boldsymbol{E}(X)$ and $\boldsymbol{E}^-(\neg X) = \{X\} = \boldsymbol{E}(\neg X)$.

The proof of completeness applies the standard technique because $\mathsf{LSO}^{\doteq}$ is essentially a first-order reasoning system. A few adjustments are needed for handling the deviations from LK. We must ensure not only that all formulas are processed and all terms are substituted by $(\forall_L)$, but also that all sentences are substituted by $(\forall_L^\phi)$.

Fact 5.3 (2.15). *If $\Gamma \nvdash \Delta$ in $\mathsf{LSO}^{\doteq}$ then there is a language graph $G(\mathcal{L}^{\doteq})$ with a semikernel L such that (i) $\Gamma \subseteq L$ and (ii) $\Delta \subseteq \boldsymbol{E}^-_G(L)$, hence L covers $\Gamma \cup \Delta$.*

Proof. Fresh s-eigenvariables introduced by $(\forall_R^\phi)$ come from a set Θ disjoint from the set of bound s-variables. (The same can be assumed for o-eigenvariables X but these are treated in the standard way and ignored below.) We assume an enumeration E_F of all formulas $\boldsymbol{F}$ (over o-eigenvariables X and with atoms $\boldsymbol{A}$ including s-eigenvariables Θ), where each formula occurs infinitely often, and an enumeration $E_T = t_1, t_2...$ of terms $\boldsymbol{T}_X$ so that each occurs infinitely often. We enumerate all triples $\langle S_i, t_j, S_k\rangle \in \boldsymbol{F} \times \boldsymbol{T} \times \boldsymbol{F}$, with each $\langle S_i, t_j, \_\rangle$ and $\langle S_i, \_, S_k\rangle$ occurring infinitely often. This is interleaved with an enumeration of all pairs $\langle S_i, S_j\rangle \in \boldsymbol{F} \times \boldsymbol{F}$, each occurring infinitely often, yielding enumeration $\boldsymbol{En}$.

1. In the usual way, we construct a derivation tree starting with the root $\Gamma \vdash \Delta$ to be proven. An *active* sequent – initially, only the root – is a non-axiomatic leaf of the tree constructed bottom-up so far. At each step, we consider the next item from $\boldsymbol{En}$. For $\langle S_i, t_j, S_k\rangle$, we find the active occurrences (i.e., occurrences in the active sequents) of S_i. If none is found, we proceed to the next element of $\boldsymbol{En}$. Pairs $\langle S_i, S_j\rangle$ serve treatment of $\doteq$ atoms.

1.i. Encountering $\langle S_i, S_j\rangle$, we apply rules for $\doteq$.

(a) If S_i and S_j are syntactically identical, we add $S_i \doteq S_i$ to the left of active sequents. Otherwise, if $S_i \doteq S_j$ occurs on the left of an active sequent

(b) along with some $P[S_i]$, we add $P[S_j]$ (similarly, we add $P[S_i]$ to an occurring $P[S_j]$), and

(c) if $S_i = P(A_1, ..., A_n)$ while $S_j = P(B_1, ..., B_n)$, we add $A_1 \doteq B_1, ..., A_n \doteq B_n$.

Case (b) steps down only to the immediate arguments of the outermost operator P.

The remaining cases address triples $\langle S_i, t_j, S_k\rangle$ encountered along $\boldsymbol{En}$.

1.ii. If $S_i \in \boldsymbol{A}$, or S_i has no active occurrences, proceed to the next item in the enumeration.

1.iii. Otherwise, retain S_i from the active sequent, which instantiates the conclusion of the relevant rule, in the new leaves obtained from the rule's premises. E.g., if $S_i = A \wedge B$ then every active sequent of the form $\Gamma', A \wedge B, \Gamma'' \vdash \Delta$ is replaced by

$$\frac{A, B, \Gamma', A \wedge B, \Gamma'' \vdash \Delta}{\Gamma', A \wedge B, \Gamma'' \vdash \Delta}$$

while every active sequent of the form $\Gamma \vdash \Delta', A \wedge B, \Delta''$ by

$$\frac{\Gamma \vdash A, \Delta', A \wedge B, \Delta'' \qquad \Gamma \vdash B, \Delta', A \wedge B, \Delta''}{\Gamma \vdash \Delta', A \wedge B, \Delta''}$$

Analogously for $S_i = \neg A$, ignoring the other elements of the triple.

1.iv. If $S_i = \forall x D(x)$, each active sequent of the form $\Gamma', \forall x D(x), \Gamma'' \vdash \Delta$, is replaced by the derivation with a new leaf adding $D(t_j)$ to its antecedent

$$\frac{D(t_j), \Gamma', \forall x D(x), \Gamma'' \vdash \Delta}{\Gamma', \forall x D(x), \Gamma'' \vdash \Delta}\ .$$

Every active sequent of the form $\Gamma \vdash \Delta', \forall x D(x), \Delta''$ is replaced by (x is a fresh eigenvariable):

$$\frac{\Gamma \vdash \Delta', \forall y D(y), \Delta'', D(x)}{\Gamma \vdash \Delta', \forall y D(y), \Delta''}$$

1.v. If $S_i = \forall \phi D(\phi)$, then replace every active sequent of the form $\Gamma', \forall \phi D(\phi), \Gamma'' \vdash \Delta$ by

$$\frac{D(S_k), \Gamma', \forall \phi D(\phi), \Gamma'' \vdash \Delta}{\Gamma', \forall \phi D(\phi), \Gamma'' \vdash \Delta}$$

while every active sequent of the form $\Gamma \vdash \Delta', \forall \phi D(\phi), \Delta''$ by

$$\frac{\Gamma \vdash \Delta', D(\alpha), \Delta''}{\Gamma \vdash \Delta', \forall \phi D(\phi), \Delta''}$$

for a fresh s-eigenvariable $\alpha \in \Theta$.

2. A branch gets closed when its leaf is an axiom, and the tree closes when all branches do. This yields a proof. Otheriwse, either some branch terminates with a non-axiomatic irreducible sequent, or the tree is obtained as the ω-limit of this process. A finite non-axiomatic branch gives easily a countermodel. We show that also an infinite branch gives a countermodel of all sequents on this branch, including the root sequent.

3. If β is an infinite branch, with β_L/β_R all formulas occurring in β on the left/right of $\vdash$, then there is a language graph G with a semikernel L such that $\beta_L \subseteq L$ and $\beta_R \subseteq \boldsymbol{E}^-(L)$. The rest of the proof establishes this claim. Absence of any (Ax) in β implies that $\beta_L \cap \beta_R = \varnothing$, which is often applied implicitly.

If β contains any FOL-atoms, construct first a FOL-structure M, giving a countermodel to $(\beta_L \cap \boldsymbol{S}_M^-) \Rightarrow (\beta_R \cap \boldsymbol{S}_M^-)$, in the standard way; otherwise, set $M = \varnothing$. Let $G = G_M(\mathcal{L}^{\doteq})$.

4. First, consider $\doteq$-atoms, denoting by Eq_L/Eq_R the equations in β_L/β_R.

4.i. Eq_L contains a unifier U of Eq_L that does not unify any equation in Eq_R.

Since each $A \doteq A$ is in Eq_L by **1.i.**(a), no such occurs on the right and each $S \doteq T \in Eq_R$ has syntactically distinct sentences. Each finite subset of Eq_L is unifiable since the branch is not (uniAx). Consequently, Eq_L is unifiable, since unifiability has finite character. A possible failure of a systematic unification process **1.i.**(b) and (c), reflecting applications (rep) and (unif), happens in a finite time (cf. Property 4 in [2]). In fact, Eq_L contains its unifier. If $P(A) \doteq P(B) \in Eq_L$ then also $A \doteq B$ by **1.i.**(c). This continues to the trivial equations $C \doteq C$, or $\phi \doteq C$, and the latter give such a unifier $U \subseteq Eq_L$.

Since $\beta_L \cap \beta_R = \varnothing$, if $S \doteq T \in Eq_R$ then S and T are not unifiable by equations from Eq_L. For suppose that they are, e.g., $S \doteq T$ is $P(\phi) \doteq P(Q)$, while equation $\phi \doteq Q$ enters Eq_L at some stage. $P(Q) \doteq P(Q)$ enters Eq_L by **1.i**.(a), and then also $P(\phi) \doteq P(Q)$ by **1.i**.(b), giving (Ax). (If $S \doteq T \in Eq_R$ is unifiable but not by any equations from Eq_L then it is not a trivial equation, i.e., S and T are syntactically distinct, with the free variables that, turning into constants in the countermodel, make $S \not\doteq T$.)

For each $(\phi \doteq S) \in U$ with $\phi \in \Theta$, we replace all occurrences of ϕ in β with S. This leaves only trivial equations $S \doteq S$ on the left, and only false ones $S \doteq T$ on the right. We continue we the β so obtained.

4.ii. Since β is over $\boldsymbol{S}_M \supseteq \boldsymbol{A}_M \supseteq \Theta$, we could add Θ as new sentential constants to the language and the graph, but this would introduce changes that can be avoided, as we now show. For each $\theta \in \Theta$ occurring in β, we choose a witness from $\boldsymbol{S}$, based on the fact that:

– neither β_L nor β_R contains all sentences, as any such sequent is provable.
– In fact, neither contains all atoms. If any does, this happens only thanks to $(\forall_L^\phi)$ instantiating $\forall\phi F(\phi) \vdash ...$ with each atom A, extracted then from $F(A)$ by other rules. In particular, ϕ must have a sentential occurrence in F. But then also every other sentence, instantiating ϕ, is extracted in the same way, so that all sentences actually appear.
– Finally, neither β_L nor β_R contains infinitely many atoms, since they could only come from the source excluded in the previous point.

In an infinite failure branch β, to each $\theta \in \Theta$ occurring on one side, assign some atom not occurring on the other. Do it repeatedly for all eigenvariables from Θ occurring on β. (We can also assign $A \vee \neg A$ to each $\in \beta_L$ and $A \wedge \neg A$ to each $\theta \in \beta_R$, for some atom A not occurring elsewhere, to which the countermodel can then assign any value.)

5. With β modified as described in **4.i** and **4.ii**, we show that
(def) $L = \beta_L \cup Z$, where $Z = (\boldsymbol{E}(\beta_R) \cap \boldsymbol{E}^-(\beta_R)) \setminus \beta_L$,
is a required semikernel of G. By 4.i, β_L contains only trivial equations, while β_R only $S \doteq T$ for syntactically distinct S, T, so L satisfies the requirement for the intended interpretation of $\doteq$.

5.i. (Def) gives $\beta_L \subseteq L$ and $\boldsymbol{E}(\beta_R) \cap \boldsymbol{E}^-(\beta_R) \subseteq L$, but we have to verify $\beta_R \subseteq \boldsymbol{E}^-(L)$. Indeed, no $X \in \beta_R \setminus \boldsymbol{E}^-(L) = \beta_R \setminus \big(\boldsymbol{E}^-(\boldsymbol{E}^-(\beta_L)) \cup \boldsymbol{E}^-(\boldsymbol{E}(\beta_R) \cap \boldsymbol{E}^-(\beta_R))\big)$, for if $X \in \beta_R$ is:

1. an atom, then $\boldsymbol{E}(X) \cap \boldsymbol{E}^-(X) = \{\neg X\} \subseteq Z \subseteq L$;
2. $\neg Y$, then $Y \in \beta_L$ and $X \in \boldsymbol{E}^-(L)$;
3. $A_1 \wedge A_2$, then some $A_i \in \beta_R$ and $\neg A_i \in \boldsymbol{E}(X) \cap \boldsymbol{E}^-(A_i) \subseteq Z \cup \beta_L$, so $X \in \boldsymbol{E}^-(L)$;
4. $\forall\phi F\phi$, then $F\psi \in \beta_R$, for some $\psi \in \Theta$, and $\neg F\psi \in \boldsymbol{E}(X) \cap \boldsymbol{E}^-(F\psi) \subseteq Z \cup \beta_L$.

6. We show $\boldsymbol{E}(L) \subseteq \boldsymbol{E}^-(L) \cap (\boldsymbol{V} \setminus L)$ separately for each part $\beta_L \cup Z = L$. The subpoints below establish $\boldsymbol{E}(\beta_L) \subseteq \boldsymbol{E}^-(L) \cap (\boldsymbol{V} \setminus L)$, considering cases of $A \in \beta_L$.

6.i. For an atom $A \in \boldsymbol{A}_M$, since $A \in \beta_L \subseteq L$ so $A \notin \beta_R$, hence $\neg A \notin \beta_L$ and, since $\boldsymbol{E}(\neg A) \overset{5.2}{=} \{A\}$, $\neg A \notin \boldsymbol{E}^-(\beta_R)$. Thus $\boldsymbol{E}(A) \overset{5.2}{=} \{\neg A\} \subseteq \boldsymbol{E}^-(A) \cap (\boldsymbol{V} \setminus L) \subseteq \boldsymbol{E}^-(L) \cap (\boldsymbol{V} \setminus L)$.

6.ii. $A = \neg C \in \beta_L$ implies $C \in \beta_R$, so $\boldsymbol{E}(A) \overset{5.2}{=} \{C\} \subseteq \beta_R \subseteq \boldsymbol{E}^-(L)$ by **5.i**.

We show $\boldsymbol{E}(A) \subset \boldsymbol{V} \setminus L$. $C \notin \beta_L$ since $\beta_L \cap \beta_R = \varnothing$. Suppose $C \in \boldsymbol{E}(\beta_R) \cap \boldsymbol{E}^-(\beta_R)$. If $C = \neg D$ then $\neg D \in \boldsymbol{E}^-(\beta_R)$, i.e., $\boldsymbol{E}(\neg D) \overset{5.2}{=} \{D\} \subset \beta_R$, while $A = \neg C = \neg\neg D \in \beta_L$ implies also $\neg D \in \beta_R$ and $D \in \beta_L$, contradicting $\beta_L \cap \beta_R = \varnothing$.

Otherwise, i.e., if C does not start with $\neg$, then for any $F \in \beta_R$ for which $C \in \boldsymbol{E}(F)$, Fact 5.2.(3-4) forces $F = \neg C = A$, contradicting $\beta_R \cap \beta_L = \varnothing$.

6.iii. $A = B \wedge C \in \beta_L$ implies $\{B, C\} \subset \beta_L$ and $\{\neg B, \neg C\} \cap \beta_L = \varnothing$, so $\boldsymbol{E}(B \wedge C) \overset{5.2}{=} \{\neg B, \neg C\} \subseteq \boldsymbol{V} \setminus \beta_L$ and $\boldsymbol{E}(B \wedge C) = \{\neg B, \neg C\} \overset{5.2}{\subseteq} \boldsymbol{E}^-(\{B, C\}) \subseteq \boldsymbol{E}^-(\beta_L)$. If, say, $\neg B \in \boldsymbol{E}^-(\beta_R)$, then $B \in \beta_R$ would contradict $\beta_L \cap \beta_R = \varnothing$. The same if $\neg C \in \boldsymbol{E}^-(\beta_R)$. Thus, $\boldsymbol{E}(B \wedge C) \subseteq \boldsymbol{E}^-(L) \cap (\boldsymbol{V} \setminus L)$.

6.iv. $A = \forall\phi D(\phi) \in \beta_L \Rightarrow \{D(S) \mid S \in \boldsymbol{S}_M\} \subseteq \beta_L$, so $\boldsymbol{E}(\forall\phi D(\phi)) \overset{5.2}{=} \{\neg D(S) \mid S \in \boldsymbol{S}_M\} \overset{5.2}{\subseteq} \boldsymbol{E}^-(\{D(S) \mid S \in \boldsymbol{S}_M\}) \subseteq \boldsymbol{E}^-(\beta_L) \subseteq \boldsymbol{E}^-(L)$.

If any $\neg D(S) \in L$ then either $\neg D(S) \in \beta_L$, so $D(S) \in \beta_R$, or $\neg D(S) \in \boldsymbol{E}(\beta_R) \cap \boldsymbol{E}^-(\beta_R)$, which implies $D(S) \in \beta_R$, since $\boldsymbol{E}(\neg D(S)) \overset{5.2}{=} \{D(S)\}$. In either case, $D(S) \in \beta_R$ contradicts $\beta_L \cap \beta_R = \varnothing$. Thus $\boldsymbol{E}(\forall \phi D(\phi)) \subseteq \boldsymbol{V} \setminus L$.

6.v. For $A = \forall x.D(x)$, the argument is as in **6.iv**. $\forall x.D(x) \in \beta_L$ implies $\{D(t) \mid t \in \boldsymbol{T}_M\} \subseteq \beta_L$, so $\boldsymbol{E}(\forall x D(x)) \overset{5.2}{=} \{\neg D(t) \mid t \in \boldsymbol{T}_M\} \overset{5.2}{\subseteq} \boldsymbol{E}^-(\{D(t) \mid t \in \boldsymbol{T}_M\}) \subseteq \boldsymbol{E}^-(\beta_L) \subseteq \boldsymbol{E}^-(L)$.

If any $\neg D(t) \in L$, then either $\neg D(t) \in \beta_L$, so $D(t) \in \beta_R$, or $\neg D(t) \in \boldsymbol{E}(\beta_R) \cap \boldsymbol{E}^-(\beta_R)$, which implies $D(t) \in \beta_R$, since $\boldsymbol{E}(\neg D(t)) \overset{5.2}{=} \{D(t)\}$. In either case, $D(t) \in \beta_R$ contradicts $\beta_L \cap \beta_R = \varnothing$. Thus $\boldsymbol{E}(\forall x.D(x)) \subseteq \boldsymbol{V} \setminus L$.

7. Also for each sentence $S \in Z = (\boldsymbol{E}(\beta_R) \cap \boldsymbol{E}^-(\beta_R)) \setminus \beta_L$ it holds that $\boldsymbol{E}(S) \subseteq \boldsymbol{E}^-(L) \cap (\boldsymbol{V} \setminus L)$:

7.i. If $S \in Z$ does not start with $\neg$, then $\boldsymbol{E}^-(S) \overset{5.2}{=} \{\neg S\}$, so $\neg S \in \beta_R$, implying $S \in \beta_L$, so $S \notin Z$.

7.ii. If $S = \neg A \in Z \subseteq \boldsymbol{E}^-(\beta_R)$ then $\boldsymbol{E}(\neg A) \overset{5.2}{=} \{A\} \subset \beta_R \overset{\mathbf{5.i}}{\subseteq} \boldsymbol{E}^-(L)$. If $A \in Z$, then by **7.i** it starts with $\neg$, i.e., $A = \neg B$ and $\boldsymbol{E}(\neg B) \overset{5.2}{=} \{B\} \subset \beta_R$. Since also $A \in \beta_R$ so $B \in \beta_L$, contradicting $\beta_L \cap \beta_R = \varnothing$. Hence $A \notin Z$ and $A \notin \beta_L$ (since $A \in \beta_R$), i.e., $A \notin L = Z \cup \beta_L$, so that $\boldsymbol{E}(\neg A) = \{A\} \subseteq \boldsymbol{V} \setminus L$.

By points **6** and **7**, $\boldsymbol{E}(L) = \boldsymbol{E}(\beta_L) \cup \boldsymbol{E}(Z) \subseteq \boldsymbol{E}^-(L) \cap (\boldsymbol{V} \setminus L)$, so $L \in SK(G)$. By **4** it contains trivial equations and negations of the false ones, thus respecting the intended interpretation of $\doteq$, and by **5.i** satisfies (i)-(ii). □

REFERENCES

[1] Walter Burleigh, ***De puritate artis logicae tractatus longior***, The Franciscan Institute St. Bonaventure, New York, 1955, [edited by Philotheus Boehner].

[2] Wan Fokkink, *Unification for infinite sets of equations between finite terms*, ***Information Processing Letters***, vol. 62 (1997), pp. 183–188.

[3] Vann McGee, *Maximal consistent sets of instances of Tarski's schema (T)*, ***Journal of Philosophical Logic***, vol. 21 (1992), pp. 235–241.

[4] Sara Negri and Jan von Plato, ***Structural proof theory***, Cambridge University Press, 2001.

[5] Victor Neumann-Lara, *Seminúcleos de una digráfica*, ***Technical report***, Anales del Instituto de Matemáticas II, Universidad Nacional Autónoma México, 1971.

[6] Douglas Patterson, *Inconsistency theory of semantic paradox*, ***Philosophy and Phenomenological Research***, vol. 79 (2009), no. 2, pp. 387–422.

[7] Michale Sheard, *A guide to truth predicates in the modern era*, this Journal, vol. 59 (1994), no. 3, pp. 1032–1054.

[8] E. A. Synan, *The universal and supposition in a Logica attributed to Richard of Campsall*, ***Nine mediaeval thinkers*** (J. R. O'Donnell, editor), Pontifical Institute of Mediaeval Studies, Toronto, 1955.

[9] Alfred Tarski, *The concept of truth in formalised languages*, ***Logic, semantics, metamathematics, papers from 1923 to 1938***, Oxford at the Clarendon Press, 1956, [translation of the Polish original from 1933 by J. H. Woodger].

[10] Willard van Orman Quine, *Two dogmas of empiricism*, ***The Philosophical Review***, vol. 60 (1951), pp. 20–43.

[11] Michal Walicki, *Paradoxes versus contradictions in a logic of sentential operators*, ***Logic and Logical Philosophy***, (2024), pp. 1–44.

[12] ———, *The structure of paradoxes in a logic of sentential operators*, ***Journal of Philosophical Logic***, vol. 53 (2024), pp. 1579–1639.

DEPARTMENT OF INFORMATICS
UNIVERSITY OF BERGEN, NORWAY
E-mail: Michal.Walicki@uib.no