# Modeling Truth for Semantics

Ori Simchen

**Abstract**

The Tarskian notion of truth-in-a-model is the paradigm formal capture of our pre-theoretical notion of truth for semantic purposes. But what exactly makes Tarski's construction so well suited for semantics is seldom discussed. In Simchen 2017 a certain requirement on the successful formal modeling of truth for semantics – "locality-per-reference" – is articulated against a background discussion of metasemantics and its relation to truth-conditional semantics. It is a requirement on any formal capture of sentential truth vis-à-vis the interpretation of singular terms and it is clearly met by the Tarskian notion. In this paper another such requirement is explored – "locality-per-application" – which is a requirement on a formal capture of sentential truth vis-à-vis the interpretation of predicates. This second requirement is also clearly met by the Tarskian notion. The two requirements taken together offer a fuller answer than has been hitherto available to the question of what makes Tarski's notion of truth-in-a-model especially well suited for semantics.

How should truth be modeled for the purposes of truth-conditional semantics? The received paradigm is Tarski's work. Given the work's prominence and centrality for subsequent semantic theorizing, it is actually easy to forget that what Tarski did was offer a certain theoretical capture of an everyday notion, the notion of sentential truth. Holding the theoretical capture apart from the everyday notion allows us to reflect on the achievement by asking what makes work on truth especially suited for semantics. The general question of how sentential truth should be modeled for semantic purposes is neither trivial nor uninteresting and yet seldom discussed in its own right. Recently, a preliminary exploration of an answer has appeared in Simchen 2017 against a background discussion of metasemantics and its relation to formal semantics. The purpose of this paper is to make further progress on the issue.

The first order of business is to explain the context in which the question of how truth should be modeled arises in previous work on the topic, particularly in metasemantics. Think of metasemantics as the study of what determines that expressions have their semantic significance. There is a fault line in metasemantics between positions that portray semantic endowment as determined directly by conditions surrounding the production or employment of the items thus endowed and positions that portray semantic endowment as determined by conditions surrounding the interpretive consumption of such items. Examples of metasemantic views of the first sort are Donnellan's views on how certain uses of descriptions refer, Kripke's views on how proper names come to name whatever they name, Kaplan's views on how demonstrative pronouns come to refer to their demonstrata, and Putnam's views on how kind terms come to apply to instances of the relevant kinds. Examples of metasemantic views of the latter sort are Davidson's interpretationist notion that expressions are assigned semantic values so as to generate the right ("interpretive") truth-conditions for sentences in context and the Lewisian interpretationist doctrine of reference magnetism according to which expressions are assigned semantic values so as to maximize truth for the total theory in which they are embedded while respecting objective joints in nature. Views of the former sort prioritize semantic endowment for sub-sentential expressions over semantic endowment for whole sentences in the order of metasemantic explanation. Views of the latter sort prioritize semantic endowment for sentences over semantic endowment for sub-sentential expressions.

Chapter 2 of Simchen 2017 contains an argument against interpretationism in metasemantics that aims to show that any such position is vulnerable to radical indeterminacy in singular reference. Whether or not this argument succeeds in its broad metasemantic aims is beyond our present concern. What is of interest to us here is the argument's deployment of an alternative construal of sentential truth. The argument considers a simple first-order extensional language $L$ that contains besides the usual logical vocabulary only constants and predicate letters of various arities. A model $m$ is understood in the usual way as $\langle M, \sigma \rangle$ where $M$ is the model's universe of discourse and $\sigma$ an interpretation function that assigns to each constant a member of $M$ and to each predicate letter of arity $n$ a subset of $M^n$. As is familiar, the standard Tarskian construction of truth-in-a-model includes the following clause for the atomic cases:

$$m \models \phi(v_1, ..., v_i, t_1, ..., t_j)^s \iff \langle s(v_1), ..., s(v_i), \sigma(t_1), ..., \sigma(t_j) \rangle \in \sigma(\phi),$$

where $s$ is an assignment function that assigns members of $M$ to the free variables $v_1, ..., v_i$, where $t_1, ..., t_j$ are constants, and where $\phi$ is an $n$-place predicate letter $(n = i + j \geq 1)$. The alternative construal of sentential truth agrees with the standard Tarskian construal in all respects except for the atomic clause. The construction begins with Lewisian reference magnetism as a token interpretationist metasemantics. (It can also be adapted to a truth-theoretic setting so as to target Davidsonian interpretationism.) Letting $m^L$ be $\langle M^L, \sigma^L \rangle$, where $M^L$ is the intended domain, it is first assumed that for any $\sigma$, $\sigma \neq \sigma^L$, $\sigma$ is no more eligible in Lewis's sense than $\sigma^L$ as an overall interpretation of the language when it comes to the predicates.[1] $\sigma^L$ is thus maximally eligible by Lewisian standards – the interpretations of predicate letters are maximally natural in Lewis's sense. Next, an alternative model $m' = \langle M^L, \sigma' \rangle$ is considered where $\sigma'(\phi) = \sigma^L(\phi)$ for every $\phi$ so that maximal naturalness for the predicates is preserved. Letting $f : M^L \longrightarrow M^L$ be a nontrivial permutation such that for some constant constant $t^*$, $f(\sigma^L(t^*)) \neq \sigma^L(t^*)$, $\sigma'(t)$ is defined as $f(\sigma^L(t))$ for every consant $t$. Observing that the same sentences need not come out true in $m^L$ and in $m'$, a new formal capture of sentential truth is articulated under which the same sentences must come out true in $m^L$ and in $m'$: *scrambled-truth-in-a-model* ($\models^\rho$). For any $m = \langle M, \sigma \rangle$ a scrambler $\rho : M \longrightarrow M$ is a permutation on $M$. The definition of $\models^\rho$ is like that of the Tarskian $\models$ except for the atomic clause:

$$m \models^\rho \phi(v_1, ..., v_i, t_1, ..., t_j)^s \iff \langle \rho(s(v_1)), ..., \rho(s(v_i)), \rho(\sigma(t_1)), ..., \rho(\sigma(t_j)) \rangle \in \sigma(\phi).$$

Scrambled-truth-in-a-model in effect generalizes the two-place Tarskian notion by adding a scrambler $\rho$ as a third parameter, just as truth-in-a-model in effect generalizes monadic truth by adding a model $m$ as a second parameter. Truth-in-a-model thus becomes a special case of scrambled-truth-in-a-model when $\rho$ is identity, just as truth becomes a special case of truth-in-a-model when $m$ is "intended". It is then easily shown that for any sentence $S$ of $L$, $m^L \models S$ iff $m' \models^{f^{-1}} S$. (The proof is omitted but follows straightforwardly from the definitions – the scrambler $f^{-1}$ "undoes" the effect of the non-trivial permutation $f$.) The metasemantic upshot is that nothing the reference magnetist can offer qua metasemantic interpretationist will decide which of $m^L$ and $m'$ is intended. $\sigma^L$ and $\sigma'$ are equally maximally eligible by Lewisian standards and can differ as radically as we like on the interpretation of $L$'s singular terms.

Let us note that this kind of indeterminacy in singular reference is not a variant on the type of indeterminacy achieved by Putnam's (1981) famous permutation argument and its ilk. According to the above construction the interpretation of predicate letters can remain fixed while the interpretation of the singular terms varies by utilizing the triadic scrambled-truth-in-a-model construction. That is why reference magnetism, despite its insistence on relative fixity in the interpretation of predicate letters, seems incapable of blocking the argument without special pleading.

[1] An interpretation is more eligible in Lewis's sense the more it respects objective joints in nature. See Lewis 1983 and 1984 for the original articulation of the idea and some of its ramifications. No claim is being made here as to Lewis's considered metasemantic position as articulated in Lewis 1975. However, the idea of tacking eligibility of interpretation in terms of naturalness onto the metasemantic account of Lewis 1975 is difficult to motivate. See Simchen 2017: Appendix I.

But now stepping back and considering scrambled-truth-in-a-model in its own right raises a question of independent interest: what advantage does Tarski's apparatus of truth-in-a-model enjoy over scrambled-truth-in-a-model as a formal capture of sentential truth for semantic purposes? – For clearly we cannot take scrambled-truth-in-a-model seriously as an alternative to truth-in-a-model for truth-conditional semantics. The qualification "for truth-conditional semantics" is important because even if some significant abstract model theoretic difference between the two alternative captures of sentential truth can be identified, this by itself does not without argument entail a decisive difference for the purposes of semantic theory.[2] Simchen's 2017 answer to the advantage question is that the Tarskian construal of sentential truth respects a natural and intuitive locality-per-reference requirement on modeling sentential truth: truth for singular sentences must be directly dependent on reference for singular terms. Tarksian truth-in-a-model clearly abides by the requirement while scrambled-truth-in-a-model clearly flouts it – truth in the atomic cases on that alternative scheme depends on reference for singular terms only as mediated by a scrambler. This turns out to be important in the metasemantic context. Interpretionism as a metasemantic orientation makes semantic endowment for singular terms beholden to semantic endowment for whole sentences. It sits ill with the requirement of locality-per-reference insofar as it allots priority to truth over reference. Non-interpretationist views, on the other hand, recognize reference as fixed prior to truth and thus seem at a clear advantage. They have necessary resources to account for the fact that scrambled-truth-in-a-model is not what we want from a formal capture of sentential truth. Semantics, it thus appears, is not metasemantically neutral.

The focus of the indeterminacy argument in Simchen 2017 is metasemantics, the study of how it is that expressions become "loaded" with their contributions to truth-conditions. But the proposed requirement of locality-per-reference on modeling truth for semantics raises a large concern of independent interest. Tarski's "straight" notion of truth-in-a-model, as compared with Simchen's "bent" notion of scrambled-truth-in-a-model, is so very clearly superior in modeling the everyday notion of sentential truth for semantic purposes. The question is what accounts for this phenomenology. Regardless of what one is inclined to think about the prospects and orientation of the metasemantic project, the comparison of Tarski's notion with neighboring notions seems of fundamental importance to the philosophy of the science of formal semantics. For given the prominence of Tarski's work for subsequent semantic theorizing, such a comparison enables us to achieve a better understanding of theoretical choices that lie at the core of truth-conditional semantics.

[2]For further discussion of this point see Simchen 2017: 47. See also further discussion as it pertains to the different construction to be discussed below.

With this in mind, let us now ask: what *other* requirements might there be on modeling truth beyond locality-per-reference? My present aim is to articulate another such requirement. To fix ideas consider again the Lewisian intended model $m^L = \langle M^L, \sigma^L \rangle$ with the maximally eligible $\sigma^L$. Let us now define an interpretation $\sigma''$, $\sigma'' \neq \sigma^L$, that agrees with $\sigma^L$ on the assignments to every constant but potentially disagrees on the assignments to the predicate letters in the following way.[3] For each $n$ for which $L$ contains a predicate letter of that arity we consider a nontrivial permutation $g_n$ on the set $\{\sigma^L(`P^n_1'), \sigma^L(`P^n_2'), \sigma^L(`P^n_3'), ...\}$ of assignments to all of $L$'s predicate letters '$P^n_1$', '$P^n_2$', '$P^n_3$',... (if such a nontrivial permutation exists, otherwise we let $g_n$ go trivial).[4] Now, for any arity $n$ and any predicate letter $\phi$ of this arity we define $g(\sigma^L(\phi)) = g_n(\sigma^L(\phi))$ and define $\sigma''(\phi) = g(\sigma^L(\phi))$ for each $\phi$ and $\sigma''(t) = \sigma^L(t)$ for each constant $t$. Where $m''$ is $\langle M^L, \sigma'' \rangle$ there is again no guarantee that the same sentences will come out true in $m^L$ and $m''$.

But now consider a different construal of truth – call it *jumbled-truth-in-a-model*. For any $m = \langle M, \sigma \rangle$ a jumbler $\tau$ is a permutation on $M^n$ for each $n$. Jumbled-truth-in-a-model ($\models_\tau$) behaves just like truth-in-a-model except for the atomic cases:

$$m \models_\tau \phi(v_1, ..., v_i, t_1, ..., t_j)^s \iff \langle s(v_1), ..., s(v_i), \sigma(t_1), ..., \sigma(t_j) \rangle \in \tau(\sigma(\phi)).$$

Truth-in-a-model becomes a special case of jumbled-truth-in-a-model when $\tau$ is identity.

Now, it follows from the definitions that for any sentence $S$ of $L$, $m^L \models S$ iff $m'' \models_{g-1} S$. We can see this by focusing on atomic sentences $\phi(t_1, ..., t_n)$ – generalizing to atomic formulas is trivial and full generality follows by induction on syntactic complexity:

$$\begin{aligned}
&m'' \models_{g-1} \phi(t_1, ..., t_n) \iff \\
&\langle \sigma''(t_1), ..., \sigma''(t_n) \rangle \in g^{-1}(\sigma''(\phi)) \iff \\
&\langle \sigma^L(t_1), ..., \sigma^L(t_n) \rangle \in g^{-1}(g(\sigma^L(\phi))) \iff \\
&\langle \sigma^L(t_1), ..., \sigma^L(t_n) \rangle \in \sigma^L(\phi) \iff \\
&m^L \models \phi(t_1, ..., t_n).
\end{aligned}$$

We note that $m^L$ and $m''$ are equally maximally eligible by Lewisian standards. So under the minimal assumption that there be an $n$ for which the language $L$ contains more than a single predicate letter of that arity, the upshot is indeterminacy in the

[3]The basic idea is due to Carl Posy.

[4]Clearly for any $n$ for which $L$ contains only a single predicate letter of that arity the requirement of nontriviality cannot be met.

application of predicates sustained by the availability of an alternative construal of sentential truth. It is yet another threat of semantic indeterminacy that passes under the radar of reference magnetism.

Jumbled-truth-in-a-model opens up yet another interesting challenge: what advantage does truth-in-a-model have over jumbled-truth-in-a-model as a formal capture of sentential truth? – For surely jumbled-truth-in-a-model cannot be taken seriously for semantic purposes despite the fact that unlike scrambled-truth-in-a-model, jumbled-truth-in-a-model does not violate the aforementioned locality-per-reference constraint. And yet all the same it is as unsuitable for modeling sentential truth for truth-conditional semantics as scrambled-truth-in-a-model.

We note that an answer to this new advantage question cannot merely point to some abstract feature truth-in-a-model has and jumbled-truth-in-a-model lacks without further argument as to why having this feature should matter to semantics. For example, truth-in-a-model exhibits invariance under isomorphism:

(IUI) $$\text{If } m \models S \text{ and } m \cong m^\bullet, \text{then } m^\bullet \models S.$$

But it is surely not the case that for any $m = \langle M, \sigma \rangle$, if $m \models_\tau S$ and $m \cong m^\bullet$, then $m^\bullet \models_\tau S$. For suppose that $m \models_\tau \phi(t_1, ..., t_n)$. Then for any $m^\bullet = \langle M^\bullet, \sigma^\bullet \rangle$ for which $M \cap M^\bullet = \varnothing$, $\tau(\sigma^\bullet(\phi))$ will be undefined, and so $m^\bullet \models_\tau \phi(t_1, ..., t_n)$ will be undefined. Jumbled-truth-in-a-model has the following feature instead:

(IUI*) $$\text{If } m \models_\tau S \text{ and } m \cong m^\bullet, \text{then } m^\bullet \models_{\tau^\bullet} S,$$

> where $\tau^\bullet = I^\bullet \circ \tau \circ I^{\bullet-1}$ and $I^\bullet : M^n \longrightarrow M^{\bullet n}$ is a mapping such that for each $o \in M$, $I^\bullet(o) = I(o)$ where $I : M \longrightarrow M^\bullet$ is the isomorphism and for each $s \subseteq M^n$, $I^\bullet(s) = \{\langle I(o_1), ..., I(o_n)\rangle \mid \langle o_1, ..., o_n\rangle \in s\}$.

We see that jumbled-truth-in-a-model has IUI* by focusing again on atomic sentences – generalizing to atomic formulas is once again trivial and full generality follows by induction on syntactic complexity. Thus,

$$\begin{aligned}
&m \models_\tau \phi(t_1, ..., t_n) \iff \\
&\langle \sigma(t_1), ..., \sigma(t_n)\rangle \in \tau(\sigma(\phi)) \iff \\
&\langle I^\bullet(\sigma(t_1)), ..., I^\bullet(\sigma(t_n)) \in I^\bullet(\tau(\sigma(\phi))) \iff \\
&\langle \sigma^\bullet(t_1), ..., \sigma^\bullet(t_n)\rangle \in I^\bullet(\tau(\sigma(\phi))).
\end{aligned}$$

On the other hand, for each $\phi$ we have $\sigma^\bullet(\phi) = I^\bullet(\sigma(\phi))$, so that $I^{\bullet-1}(\sigma^\bullet(\phi)) = \sigma(\phi)$. Substituting in the last clause gets us:

$$\langle \sigma^\bullet(t_1), ..., \sigma^\bullet(t_n)\rangle \in I^\bullet(\tau(I^{\bullet-1}(\sigma^\bullet(\phi)))).$$

We see that $I^\bullet \circ \tau \circ I^{\bullet -1}$ is a jumbler on $M^\bullet$, from which we conclude that

$$m^\bullet \models_{\tau^\bullet} \phi(t_1, ..., t_n),$$

where $\tau^\bullet = I^\bullet \circ \tau \circ I^{\bullet -1}$.

But why exactly IUI should be important for semantic purposes – as opposed to IUI*, say – is a question that must be faced by anyone who wishes to tackle the contrast between truth-in-a-model and neighboring notions such as jumbled-truth-in-a-model "in the abstract", as it were. It is a difficult question. An answer would seem to require, at a minimum, an exploration of the scope of semantic theory in relation to the logicality of its fundamental notions. However, there is a far more obvious and direct route to why jumbled-truth-in-a-model is unsuitable for semantic purposes: jumbled-truth-in-a-model fails to respect a natural and intuitive locality-per-application requirement on modeling sentential truth that truth-in-a-model clearly respects. Locality-per-application is the requirement that sentential truth for singular sentences should depend directly on the application of the predicates. In jumbled-truth-in-a-model this is clearly flouted: for an atomic sentence to be jumbledly true in a model is not for the predicate to apply to the referents of the singular terms but for its jumbling to apply. This is obviously not so for Tarksi's original construction of truth-in-model, where truth for atomic cases depends directly on the application of the predicates.

As a side note, locality-per-application is anathema to interpretationism just as much as locality-per-reference. The interpretationist orientation to metasemantics renders semantic endowment for sub-sentential expressions beholden to the semantic endowment for the sentences in which they partake. For non-interpretationist views, on the other hand, the situation is reversed. The non-interpretationist can and must insist not only on reference for singular terms as settled prior to truth but also on predicate applicability as settled prior to truth. Once again, non-interpretationism is at a clear advantage over interpretationism as a metasemantic orientation to accompany truth-conditional semantic theory. It is the obvious metasemantic partner to the two requirements on modeling sentential truth we have been considering, locality-per-reference and locality-per-application.

On one level it is hardly surprising that scrambled-truth-in-a-model and jumbled-truth-in-a-model should be inferior to standard truth-in-a-model as theoretical captures of sentential truth. The requirements of locality-per-reference and locality-per-application are, after all, natural and intuitive. They are to be respected by any theoretical articulation of our everyday notion of truth. Insofar as semantics concerns itself with the formal modeling of language-world relations – a widespread even if not universally shared conception – it is hardly surprising that truth-in-a-model should be found suitable for semantics, so suitable in fact that we tend to overlook the features that render it so. It is perhaps more surprising that semantics as it is widely understood is difficult to reconcile with interpretationism as a metasemantic orientation. Teasing out further lessons from this last observation, especially lessons for the history of formal semantics and the seminal contributions made to it by leading metasemantic interpretationists, is a larger project for another day. In the meantime we note that comparing the Tarskian notion of truth-in-a-model with neighboring notions affords us a better understanding than hitherto available of pre-theoretical requirements that shape theoretical choices at the basis of contemporary semantics.[5]

# References


Lewis, David. (1975) 'Languages and Language', in Gunderson, Keith (ed.), *Minnesota Studies in the Philosophy of Science* VII (Minneapolis: University of Minnesota Press): 3-35.

Lewis, David. (1983) 'New Work for a Theory of Universals', *Australasian Journal of Philosophy* 61: 343-77.

Lewis, David. (1984) 'Putnam's Paradox', *Australasian Journal of Philosophy* 62: 221-236.

Putnam, Hilary. (1981) *Reason, Truth and History* (Cambridge: Cambridge UP).

Simchen, Ori. (2017) *Semantics, Metasemantics, Aboutness* (Oxford: Oxford UP).



[5]This paper emerged from a fruitful exchange with Carl Posy about the main argument of Chapter 2 of Simchen 2017. Errors are mine alone.