Class #23 - Tarski's Theory of Truth
Tarski, "The Semantic Conception of Truth and the Foundations of Semantics"

I. From Intensions to Extensions

Tarski's important article is called "The Semantic Conception of Truth and the Foundations of Semantics."
'Semantics' is a broad term, covering at least both sense and reference.
Frege argued that a proper analysis of language mut include theories of both intensions (sense or meaning) and extensions (reference).
His two-tiered approach solved three puzzles: the problem of cognitive content, the problem of empty reference, and the problem of opaque contexts.
It explains the synonymies of 'snow is white' and 'la nieve es blanca'; and 'bachelor' and 'unmarried male'.
It explains the differences between 'creature with a heart' and 'creature with a kidney' and between '2+2' and '7-3', terms with the same reference (extensions) but different meanings (senses).
But Frege's semantics is saddled with an inflated ontology of third-realm entities: propositions, concepts, and senses.

The logical empiricists tried to avoid intensions by explaining meaning in terms of observation and verification.
But without an analytic/synthetic distinction, their claim that individual sentences are confirmed and disconfirmed independently of each other, that they have meanings independent of a broader background theory, seems indefensible.
Holism killed the logical empiricist's project.

Quine urged that we could maintain meaningfulness without meanings.
But Quine's semantic holism seems incompatible with the basic tenets of compositionality.
Semantic holism says that the basic unit of meaning is an entire theory; the parts lack meaning on their own.
Compositionality says that the meaning of a complex sentence is built from the meanings of its parts.
Compositionality of meaning seems like a fundamental fact about language.
Quine's ontological relativity undermined the theory of reference which seems much more secure than the theory of meaning.
The arguments against the analytic/synthetic distinction are not always taken to be solid.
And the related arguments from the indeterminacy of translation have never been fully accepted.

Grice developed the IBS program, hoping to restore legitimacy to the theory of meaning, again without appeal to Fregean intensions, by explaining meaning facts in terms of speaker intentions.
But IBS ran into the bloated ontology of propositions that saddled Frege's semantics.
Grice explains sentence meaning in terms of speaker meaning, and speaker meaning in terms of belief states.
But belief states are best understood as relations between people and propositions.
The IBS account appears unavoidably circular.
Schiffer's later work despairs of developing semantic theory in any guise.

One lesson we might take from Quine's work, and Wittgenstein's, is to abandon all hope of an intensional semantic theory.
We can just construct theories of reference and truth and see if there are any semantic facts left over.
This is Davidson's approach, relying on and repurposing Tarski's theory of truth.
Davidson urges us to put aside the thorny questions about theories of meaning and look at theories of truth and reference.

The pressure to construct satisfying theories of reference and truth is greater if they are going to explain all explicable semantic properties in the absence of a theory of meaning.
Tarski did not construct his theory of truth as a semantic theory in Frege's sense.
But Davidson believes that it can work as one.
We will look at Tarski's theory first and Davidson's use of it in the next class.

II. Formal and Ordinary Conceptions of Truth

Tarski presents his theory of truth as a modest step in the direction of rehabilitating the extensionalist side of semantics.

> It is perhaps worthwhile saying that semantics...is a sober and modest discipline which has no pretensions of being a universal patent-medicine for all the ills and diseases of mankind, whether imaginary or real (345).

While the essay we are reading is fairly non-technical, Tarski's great achievement is in the technical foundations of semantics and model theory.
He constructs a formal definition of truth for artificial languages like those of logic and mathematics.
Frege and Russell dreamed of logically perfect languages, decrying the inadequacies of natural language.
Tarski is less doctrinaire about his formal work.
He does mention the hope of replacing natural languages in science with formal languages.

> At the present time the only languages with a specified structure are the formalized languages of various systems of deductive logic, possibly enriched by the introduction of certain non-logical terms. However, the field of application of these languages is rather comprehensive; we are able, theoretically, to develop in them various branches of science, for instance, mathematics and theoretical physics (347).

But, Tarski holds that the formal construction is not merely formal.

> The desired definition does not aim to specify the meaning of a familiar word used to denote a novel notion; on the contrary, it aims to catch hold of the actual meaning of an old notion (341).

> The relation between theoretical and descriptive semantics is analogous to that between pure and applied mathematics, or perhaps to that between theoretical and empirical physics; the role of formalized languages in semantics can be roughly compared to that of isolated systems in physics (365).

Tarski thus claims that he provides a precise definition of truth for formal theories and that this formal definition may help us understand our ordinary conception of truth.

Tarski models his procedure and hope after the success of formalization for contentious concepts like those of negative and imaginary numbers.
Imaginary numbers were once called impossible.
Formal theories of complex numbers assuaged worries about their legitimacy.
Similarly, Tarski hoped, formal truth theories can legitimate uses of truth in empirical science, mathematics, and even the humanities, by giving a rigorous expression of our ordinary conception.

III. Convention T

Tarski's theory of truth centers on what we call Convention T, or the T-schema.

CT   p is true if and only if x

In CT, 'p' is the name of any sentence, and x are the truth conditions of that sentence.
We can use CT to specify the truth conditions for any sentence.
Here are some instances of the T-schema.

CT1   'The cat is on the mat' is true if and only if the cat is on the mat.
CT2   '2+2=4' is true if and only if 2+2=4
CT3   'Barack Obama is president' is true if and only if the husband of Michelle Obama and father of Sasha Obama and Malia Obama is head of the executive branch of the United States of America.

Note that, as in CT3, the truth conditions on the right of the 'if and only if' need not be expressed in the same terms as the sentence on the left.
We can even use a different language for the sentence and for its truth conditions, as in CT4.

CT4   'El gato está en el alfombrilla' is true if and only if the cat is on the mat.

Notice that you could, in principle, understand the truth conditions of CT4 without understanding the meanings of the words in Spanish sentence on the left side.
That fact is important to keep in mind.

Tarski argues that the T-schema satisfies a necessary condition on truth, a condition for material adequacy.

> Now at last we are able to put into a precise form the conditions under which we will consider the usage and the definition of the term "*true*" as adequate from the material point of view: we wish to use the term "*true*" in such a way that all equivalences of the form (T) can be asserted, and *we shall call a definition of truth "adequate" if all these equivalences follow from it* (344).

Understanding how, precisely, and why the T-schema is to be used requires some work.
To see how something as simple and seemingly-obvious as CT can be important, even revolutionary, we have to put it in its intellectual and historical context.
In particular, we have to understand the profoundly unsettling effect of the semantic paradoxes on philosophy.

IV. Semantic Paradoxes, or Why All the Fuss About Truth?

In the early twentieth century, truth had taken a bad beating.
The central problem is that regimenting truth within formal theories creates paradoxes.
A paradox is a set of sentences that we have good reasons both to believe and to believe are unacceptable, or incompatible.
A sentence that appears, impossibly, both true and false is a paradox.
Various semantic notions in addition to truth create paradoxes as well.

The most important semantic paradox is the liar.

> L   This sentence is false.

L, like any sentence, could be either true or false.
If L is true, then its content, which says that L is false, is true.
So L is false.
Since L is false, the denial of its content must be true.
So, L must be not false, or true.
Thus we have good reasons to believe that L is both true and false, which is a paradox.

Paradoxes lead to contradictions.
Contradictions are unacceptable in traditional, or classical, formal systems because a contradiction entails anything.
This property of classical systems is called explosion.

Explosion	1. P • ~P	/ Q
	2. P	1, Simp
	3. P ∨ Q	2, Add
	4. ~P • P	1, Com
	5. ~P	4, Simp
	6. Q	3, 5, DS
QED		

To see how the liar leads to a contradiction, consider L again.
Applying the T-schema yields CTL.

> CTL   L is true if and only if L is false.

We can translate this sentence into first-order, Fregean logic by taking a constant, say 'p', to stand for the sentence L, and introducing a truth predicate, 'Tx'.
We take 'P is true' to be the negation of 'P is false'.

1. Tp ≡ ~Tp	From CT and the definition of 'p'
2. (Tp ⊃ ~Tp) • (~Tp ⊃ Tp)	1, Equiv
3. (~Tp ⊃ Tp) • (Tp ⊃ ~Tp)	2, Com
4. ~Tp ⊃ Tp	3, Simp
5. ~~Tp ∨ Tp	4, Impl
6. Tp ∨ Tp	5, DN

7. Tp	6, Taut
8. Tp ⊃ ~Tp	2, Simp
9. ~Tp ∨ ~Tp	8, Impl
10. ~Tp	9, Taut
11. Tp • ~Tp	7, 10, Conj

Uh-oh!

Thus, if our language contains a truth predicate, we can construct L.
If we can construct L, we can derive any and every other sentence of the language, by explosion.
But we don't really believe that every sentence is true.
Something is wrong somewhere.

Quine, in his essay "The Ways of Paradox," argues that there are grounds to question either the paradoxicality or the well-formedness of L.
It is not clear what 'this sentence' refers to.
If we substitute 'this sentence is false' for 'this sentence', then we get LQ.

LQ   'This sentence is false' is false.

LQ does not ascribe falsity to itself, and the paradox is avoided, or at least delayed.
Still, we can find other, similarly troublesome sentences.
Quine constructed QP, which avoids the above problem.

QP   'Yields falsehood when appended to its own quotation' yields falsehood when appended to its own quotation.

The problem with many sentences like L and QP seems to be rooted in the presence of semantic terms, like 'true' and 'false'.
Thus, such problematic sentences are called semantic paradoxes.

One diagnosis of many semantic paradoxes, including the liar, is that they involve illicit self-reference.
Another self-referential paradox, the barber paradox, is due to Bertrand Russell, though he credits an anonymous source.
Consider the barber in a town who shaves all the men who do not shave themselves.
Does he shave himself?
You can construct a puzzling declarative sentence, similar to the liar, which I leave to you as an exercise.

Not all semantic paradoxes involve the truth predicate.
Tarski mentions paradoxes due to Grelling and Richard.
For Grelling's paradox, note that some predicates apply to themselves, whereas others do not.
'Polysyllabic' is polysyllabic; 'monosyllabic' is not monosyllabic.
Call a predicate heterological if it does not apply to itself.
'Monosyllabic' is heterological.
'Polysyllabic' is not heterological; we can call it autological, or homological.
Now, consider whether 'heterological' applies to itself.
If it does, then 'heterological' is not heterological.
But, if 'heterological' is not heterological, then it does not apply to itself, which means that it is heterological.

We can construct a statement involving 'heterological' whose truth value is puzzling.

HH   'Heterological' is heterological.

Grelling's paradox is semantic, but does not involve 'truth' or 'falsity' explicitly.
Grelling's paradox is about meaning.

Richard's paradox, which Tarski also mentions, concerns definability and is more easily seen in a simplified form due to Berry.
Consider what number BP refers to.

BP   the least natural number not definable in fewer than twenty-three syllables

There are only finitely many expressions with fewer than twenty-three syllables.
So some natural numbers will not be definable with fewer than twenty-three syllables.
There must be a least number in this set.
But since BP has only twenty-two syllables, whatever natural number it defines is definable in fewer than twenty-three syllables.
Tilt.

One response to the semantic paradoxes is to banish semantic terms which cause them.
Tarski's theory of truth avoids the paradoxes and rehabilitates the old Aristotelian view of truth by proscribing self-reference.
He diagnoses the problem which leads to paradox as the semantic closure of natural languages.

> We have implicitly assumed that the language in which the antinomy is constructed contains, in addition to its expressions, also the names of these expressions, as well as semantic terms such as the term "*true*" referring to sentences of this language; we have also assumed that all sentences which determine the adequate usage of this term can be asserted in the language. A language with these properties will be called "*semantically closed*" (348).

If we abandon semantic closure for a formal language, we can block the semantic paradoxes.

V. Constructing a Non-Paradoxical Truth Predicate

The core of Tarski's solution is his distinction between object languages and metalanguages.
Imagine a language that does not contain the word 'true'.
Call that language the object language.
Technically, the object language contains only atomic sentences and the first-order logical structure.
Sentential functions build up all of the infinite sentences of the language.
We can not devise any of the semantic paradoxes in the object language because it is not semantically closed: it does not contain names of the terms of the object language, and it contains no truth predicate.

To examine and assess the object language, to construct a truth predicate for the object language, we ascend to a metalanguage.
We translate the object language into the metalanguage, introducing names for all of the sentences of the object language.

Then, we add a predicate, 'is true', which applies to the desired (privileged) sentences of the object language.
We thus partition the sentences of the object language into two classes: the true and the false, by invoking CT within the metalanguage.
All the theorems which follow from CT are thus sentences in the metalanguage defining the concept of truth for sentences of the object language.
'True' is a metalinguistic term applied to the object language.
The extension of a predicate is the set (or class) of objects of which that predicate holds.
The anti-extension of a predicate is the set (or class) of objects of which the predicate does not hold.
So, the extension of 'woodchuck' is the set of all woodchucks.
The anti-extension of 'woodchuck' is the class of all things that are not woodchucks.
(The anti-extension of 'woodchuck' is a class because it is too big to be a proper set!)

The extension of the truth predicate is the set of all the true sentences of the object language.
All the false sentences are in the anti-extension of the truth predicate.
Technically, we use the notion of satisfaction of an open sentence by a sequence to generate, recursively, the true sentences of the object language within the metalanguage.
We need not engage these details here; see §11.

There will be many more sentences of the meta-language than of the object language, including the theorems which follow from CT.
But the truth predicate will only apply to the sentences of the object language.
That way, we will not be able to generate the semantic paradoxes within the metalanguage.
We can not write L in the object language, because the object language does not contain the term 'false'.
And, we can not write it in the metalanguage, because while the metalanguage contains 'false', 'true' and 'false' only apply to sentences of the object language.
If we tried to construct L in the object language, 'this sentence' would refer to a sentence of the metalanguage, to which 'false' does not apply.

It may help to think of the object language and metalanguages as two different languages.
Take, for example, a fragment of Spanish which does not contain 'verdad' or any other semantic term, as your object language.
Take English as your metalanguage.
In Tarski's construction, 'true' and 'false' are English terms which apply only to Spanish sentences, and only in the metalanguage.

So, we can not write LS in the object language because it contains English and so is not a proper object-language sentence.

       LS   Este frase es false.

LS is neither Spanish nor English; it's Spanglish.
We can not write L itself in the object language either, because 'false', as we are constructing it, applies only to Spanish sentences and does not apply to English sentences like L.

If we wanted to do so, we could construct a truth predicate for the metalanguage.
But, the truth predicate for the metalanguage can not be written within the metalanguage, on pain of paradox.

So, we have to go up one step to a meta-metalanguage.
In our example, the meta-metalanguage might be Norwegian.
We could introduce a term in the meta-metalanguage, 'sann', to apply only to true English sentences.
We could introduce a term, 'uetke', to apply only to false English sentences.
Thus, Tarski's construction can lead to an infinite series of metalanguages, each one containing truth predicates only for languages lower in the hierarchy.

To determine which sentences of the object language are true and which are false, we have to examine the truth conditions given on the right hand side of instances of the T-schema.
If using a different language for the object language and the metalanguage helped, consider CT4 again.

> CT4   'El gato está en el alfombrilla' is true iff the cat is on the mat.

Remember that we can understand the truth conditions of CT4, on the right side, without understanding the meaning of the words in the Spanish sentence, on the left side.
The T-schema, CT, used in the metalanguage to define truth in the object language, is the centerpiece of Tarski's semantic conception of truth.

It is slightly misleading to call Tarski's theory a semantic conception of truth.
Such a use of 'semantic' is independent of the intensionalist meaning theories we have examined.
Any interpretation of Tarski as concerned with intensions is probably a mis-interpretation.
Truth and falsity are, strictly speaking, terms from the theory of reference.

VI. Inflationary and Deflationary Conceptions of Truth

To evaluate Tarski's so-called semantic conception of truth, we might ask what our ordinary concept of truth is and whether Tarski's construction captures it.
It is not clear what our ordinary conception of truth is.
I take our ordinary conception to be a correspondence theory, as it is traditional to interpret Aristotle.

> To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, and of what is not that it is not, is true (*Metaphysics*, 1011b25).

According to the correspondence theory of truth, truth is a relation between words and the world.
The correspondence theory is an inflationary theory of truth.
Tarski discusses other inflationary theories of truth, perhaps other versions of the correspondence theory, in §3.

> The truth of a sentence consists in its agreement with (or correspondence to) reality (343).

> A sentence is true if it designates an existing state of affairs (ibid).

All three of these inflationary characterizations of truth involve a words-worlds connection.

One worry about the correspondence theory is that we do not seem to have any extra-linguistic way to apprehend reality.
If I want to compare, say, an elephant to a picture of an elephant, or a picture of a sculpture of an elephant to a picture of an elephant, I can hold both of them in front of me, gazing from the one to the other.

If I want to compare my words to the world, I have to apprehend, on the one side, what the words mean, and on the other, the world.
But, it has seemed to some philosophers, I only apprehend the world mediately, through my ideas of it.
I do not have any access to the world as it is in itself.
(Those of you who have worked through the epistemology of the modern era will have a good understanding of the problem here.)
It seems as if I am unable to compare my words, or my ideas, to an independent world, to decide whether there really is a correspondence between them.
The correspondence theory says that truth is a matching of words to the world, but I can only really know about one side of the equation.

We can distinguish the inflationary correspondence theory from a deflationary theory of truth.
The deflationary theory of truth has many proponents, all of whom have different ways of understanding and explaining deflationism.
Common to all deflationary accounts is the belief that there is no essence to truth, no single reduction of truth to a specific property like correspondence to reality.
Some truths are linguistic.
Some are mathematical.
Some truths are conventional.
Some are empirical.
For the deflationist, the only commonality among these various kinds of truths is our use of the word 'true' to characterize them.

Some deflationists claim that truth is just a device for simplifying long conjunctions.
If you said a lot of smart things at the party, I could list them all.
Or, I could just assert LN.

LN   Everything you said last night was true.

In LN, 'true' is eliminable by a longer, clunky sentence, or set of sentences.
Such eliminations are essential to the deflationary conception of truth.

Sometimes, the deflationary theory is called a redundancy theory: to say that snow is white is true is just to say, redundantly, that snow is white.
We do not really need 'true'; it just comes in handy sometimes.

Here is a linguistic argument for the correspondence view over deflationism.
We can remove 'it is true that' from SW and leave a grammatical sentence, SW', with, arguably, the same meaning.

> SW   It is true that snow is white.
> SW'   Snow is white.

In contrast, if we remove 'is true' from OS, we are left with the ungrammatical fragment OS'

> OS   Everything Obama says is true.
> OS'   Everything Obama says

The deflationist says that all uses of truth are like, or are reducible to, the uses in SW, where truth acts as a connective, a (redundant) sentential operator.
But in OS, 'true' works like a predicate, not a connective.
The correspondence theorist takes truth to be a genuine predicate.

Tarski argues that we need a concept of truth by considering PW.

> PW   The first sentence written by Plato is true.

In PW, 'true' appears ineliminable.
More importantly, Tarski emphasizes the utility, to mathematics especially, of the claim TC.

> TC   All consequences of true sentences are true.

TC is fundamental to metalogic and model theory, fields that Tarski more or less created.
Once we start constructing formal theories, we want to know about their various properties and relations.
What theorems can or can not be proved from a given theory?
Which theories are stronger than which?
Which theories are equivalent?
In metalogic and model theory, we explore questions of whether a formal system is sound, or complete, or decidable.

Tarski uses PW and TC to show that 'truth' plays an essential role in a theory.
Philosophers have argued whether Tarski's theory is inflationary or deflationary.
Tarski's claim that 'truth' is essential may not have inflationary implications.
If 'true' is a device used to refer to other sentences, as in PW and TC, the answer to the question whether Tarski's theory can count as deflationary depends on what we think of those other sentences, the ones without 'true' and with content.
If we need a words-worlds relation in order to ascribe 'true' to a sentence, then truth will not be merely deflationary, or redundant.
If all there is to truth is eliminable, then perhaps there is no essence to truth.
Even Tarski's claim that the semantic conception of truth captures Aristotle's intent (§17) may not have

inflationary implications, since Aristotle's original claim could be given a deflationary interpretation itself!
Both inflationists and deflationists agree that a minimal condition for truth is Tarski's CT.

Inflationists and deflationists disagree about whether CT is all there is to know about truth.
The inflationist believes that there are explanations of the concept of truth inherent in the truth conditions on the right side of CT.
For the correspondence theorist, 'the cat is on the mat' is true because there is a cat, which corresponds to 'the cat', and there is a mat, which corresponds to 'the mat', and there is a relation, being on, which the cat and the mat satisfy, or in which they stand.
All other instances of the T-schema will have similar explanations in terms of the correspondence of words to worlds.

The deflationist, in contrast, believes that the T-schema is all there is to know about truth, and that there is no single kind of explanation of why all sentences are true.
'Truth' varies in application.
The explanation of the truth of CT1, for example, must differ significantly from the explanation of the truth of CT2.
Our justification for asserting statements about cats and mats relies, often, on direct observation.
Our justification for asserting statements about mathematical objects relies only indirectly (at most) on sense experience.
To repeat, according to the deflationist, we do not even need 'true' in our language. It's just a handy tool.
Deflationists look at CT as a satisfactory definition of truth.
That's why deflationism also goes by the name 'redundancy theory'.
Inflationists about truth look at CT as merely a minimal condition for truth.
They claim that there are additional requirements, like correspondence to reality.

Both inflationists and deflationists have used Tarski's work for their own purposes, and Tarski himself does not clearly distinguish between the inflationary and deflationary interpretations.
He believes that the technical merits of his semantic conception will lead us to accept it.

> It seems to me obvious that the only rational approach to [questions about the correct notion of truth] would be the following: We should reconcile ourselves with the fact that we are confronted, not with one concept, but with several different concepts which are denoted by one word; we should try to make these concepts as clear as possible (by means of definition, or of an axiomatic procedure, or in some other way); to avoid further confusions, we should agree to use different terms for different concepts; and then we may proceed to a quiet and systematic study of all concepts involved, which will exhibit their main properties and mutual relations (355).

Furthermore, Tarski believes that the semantic conception is agnostic among any deeper philosophical debates.

> We may accept the semantic conception of truth without giving up any epistemological attitude we may have had; we may remain naive realists, critical realists or idealists, empiricists or metaphysicians - whatever we were before. The semantic conception is completely neutral toward all these issues (362).

VII. Is Tarski a Deflationist?

Tarski defends calling his theory of truth a semantic conception.

> Semantics is a discipline which, speaking loosely, *deals with certain relations between expressions of a language and the objects* (or "states of affairs") "*referred to*" by those expressions. As typical examples of semantic concepts we may mention the concepts of *designation, satisfaction*, and *definition*... (345).

Hartry Field, in his famous paper on Tarski's theory, argues that Tarski is not the deflationist that the deflationists sometimes claim he is.
In order to use the T-schema, we need to supplement it with some kind of account of why we chose certain sentences to be true and not others.
To see the problem, remember that we could understand the truth conditions in CT4 without understanding the meanings of the words in the Spanish sentence on the left.
To understand the language, it seems, it is not enough just to construct the extension and anti-extension of 'is true'.
We want to understand the component parts of the Spanish expressions, and how they interact to form true or false sentences.
Convention T, by itself, does not provide that kind of explanation.
Thus, Tarski's construction only reduces 'truth' to other semantic notions, like satisfaction, designation, and reference.
It doesn't get rid of semantic notions.

Note that Field's criticism is not the criticism that Tarski considers in §15, that his concept of truth presumes, in circular fashion, logical connectives that are defined by the truth tables.
Against that (potentially devastating) objection, Tarski rightly argues that we can define all the sentential connectives together, implicitly, by their functions.
We need not appeal to the truth tables.
We can even introduce the connectives as undefined terms.
Such implicit definitions are acceptable to avoid circularity in the case of logical connectives, though they are not acceptable for terms with greater content, as, say, a way of avoiding the circularity in Quine's hocus-pocus argument.
[An explanation of that last sentence could be a good paper.
Tarski's claim in §15 about the role of the connectives in proof theory, as opposed to semantic (or model) theory is a hint.]

If we are merely concerned with constructing a metalinguistic truth predicate, Convention T and its supplements (designation, satisfaction, definition) might suffice.
We might, in contrast, wish to take Tarski's claim to constructing a semantic notion of truth seriously.
In that case, it would be useful to be able not merely to explain truth in terms of other semantic notions, but to reduce them to physical ones.
Recall the IBS attempt to reduce the semantic to the psychological, with an aim to physicalistic theories of mind.
If we want to avoid appeals to Fregean third-realm entities in explaining truth, we need to supplement Tarski's theory with something like causal accounts of reference.
We need, in addition to Convention T, an explanation of why the terms are true of the things of which they are true, in a way that is consistent with physics.

It is not that we could not add such an account to complete Tarski's theory.
But, once we do, the theory does not appear deflationary.
The concepts of satisfaction and denotation are referential, involving inflationary concepts of a words-worlds connection.

VIII. Kripke and the Truth Hierarchy

Field's worry about Tarski's theory concerns its use by deflationists in a broader semantic project.
But, there are technical concerns about his constructions, as well.
Tarski's construction produces a hierarchy of languages.
To construct a truth predicate for an object language, we eliminate semantic terms from our object language and ascend to a metalanguage.
If we want to know about the truth of sentences of our object language, we take an external perspective, working in a metalanguage.

We might reasonably wonder about truth in the metalanguage.
Of course, for the same reasons that the object language can not contain a truth predicate, the metalanguage can not contain its own truth predicate.
But, we can construct a truth predicate for the metalanguage in a further metalanguage.
To construct a truth predicate for the second metalanguage, we can construct a third, and so on.
Each of the separate truth predicates occurs at a different level in this ever-expanding hierarchy.
The relations among these truth predicates are merely analogical.
Each metalanguage is distinct, and has different terms.
Each truth predicate is independent of each other.
We are burdened not with one 'truth', but an infinite hierarchy of 'truth's.

To make matters worse, there are cases in which we do not know which level in the hierarchy any particular use of 'true' or 'false' belongs to. Consider GB.

> GB   Everything George W. Bush says is false.

GB must be made in a metalanguage one step higher than anything that Bush ever said.
If I assert GB, then to know what level my 'false' belongs to, I need to know about all the levels of Bush's uses of 'true' and 'false'.
If Bush once claimed BC, for example, then in order to know what level Bush's 'false' occurs at, we also need to know all the levels of the uses of 'true' and 'false' in whatever Clinton said.

> BC   Everything Bill Clinton says is false.

Furthermore, if Clinton were the speaker of GB, then Bush and Clinton become embroiled in a semantic circle.
The level of GB must be higher than the level of BC.
The level of BC must be higher than that of GB.
Tarski's hierarchical approach seems to lead to a contradiction in this case despite the fact that there seems to be nothing contradictory about the conjunction of GB (stated by Clinton) and BC (stated by Bush).
They are just both false statements: both Clinton and Bush have uttered some truths.

Saul Kripke, in a paper called "Outline of a Theory of Truth," showed that we can construct a truth predicate for a language embedded within the object language itself, without creating paradox.
Here is a quick sketch of Kripke's approach.
We start with a base language, containing no logical connectives, quantifiers, or truth predicate.
Then, we add a truth predicate to the language itself.
We can more-or-less easily decide which sentences of the base language are true and which are false, since there is no truth predicate in the base language.
Then, we can add the familiar logical connectives: negation, conjunction, and disjunction, say.
The semantics for the propositional connectives are easily presented, as well.
So, we can apply the truth predicate to all base-level sentences and logical functions of them.

Next, we can consider sentences with single uses of semantic terms licensed so far.
We repeat the original process, adding more complex sentences to our lists of true and false sentences.
We can proceed to sentences of greater and greater semantic complexity.
At each level, we bring along all the earlier sentences, and apply the truth predicate to them.
But, the truth predicate does not apply to sentences at its own level.
Eventually, we can, in principle, reach any of a variety of fixed points past which further construction is unwarranted.
There are many different fixed points, and lots of technical work can be done with them.

The importance of Kripke's construction is that he produces a single, object-level truth predicate, rather than a series of hierarchical languages all with one truth predicate.
This object-language truth predicate allows us to value many sentences that include 'true'.
We can have a language with all of those sentences, and one truth predicate for all of them.
Kripke contains the entire Tarskian hierarchy in one language.

IX. Summary

There are at least two ways to look at Tarski's semantic theory of truth.
The first way is minimalist, and it focuses on the condition of adequacy, the T-schema.
The second way is inflationist, and it focuses on the extent to which Tarski legitimizes our ordinary, correspondence notion of truth.
There is no question that the notion of truth is useful, in sentences like LN, and essential to metalogical work, in sentences like TC.
Tarski, and those following him, have vindicated formal theories of truth insofar as they allow us to capture these minimal uses of the term.

Deflationists often seem to be arguing against a more substantial notion of truth than correspondence.
They deny that there is an essence to truth.
But, correspondence theorists are not ordinarily committed to any kind of spooky essence.
Tarski himself makes that pretty clear.

> In no interpretation of the term 'metaphysical' which is familiar and more or less intelligible to me does semantics involve any metaphysical elements peculiar to itself (364).

I wonder if the deflationist is arguing against a straw inflationist.
Still, the question of whether philosophers need any notion of truth (or Truth) is much discussed.

Tarski makes a compelling argument that science aims at truth.
The main reason we want consistent theories is because we know that an inconsistent theory contains a falsehood.

> As soon as we succeed in showing that an empirical theory contains (or implies) false sentences, it cannot be any longer considered acceptable (367).

There are obvious epistemic worries about our access to truth, our ability to know what is true and what is false.
The old problem of whether we can assess a words-worlds connection still resonates.
Some philosophers continue to try to replace truth with a weaker condition like warranted assertability or coherence
I'm a big fan of truth, and Truth.
But, there are obvious epistemic worries about our access to it.
The old problem of whether we can assess a words-worlds connection, being embedded in one side of it, still resonates.

Whether or not Tarski's solution to the problem of the paradoxes is ideal, the distinction between object language and metalanguage has become fundamental in all contemporary formal treatments of language.
The importance of Tarski's theory of truth for us at the moment resides in Davidson's hopes for an extensionalist semantic theory, which we will examine next.