The only entirely natural theory of truth is given by Tarski's schema

\[(T) \quad Tr(T\phi) \leftrightarrow \phi\]

where \(T\phi\) is a structurally descriptive name — for our purposes, a Gödel number — of a sentence \(\phi\) of a language \(L\) and \(Tr\) is a predicate which is intended to represent the set of true sentences of \(L\). However, the Tarski–Epimenides paradox\(^2\) shows that, assuming the predicate \(Tr\) is part of the language \(L\), this natural account will be inconsistent with the basic laws of syntax.\(^3\) For some purposes, a sufficient response will be simply to insist\(^4\) that the theory of truth for a language must never be developed within the language itself, but rather within an essentially richer metalanguage. But for other purposes, one would like to be able to develop a semantic theory of a language within that very language; most urgently, one would like to be able to talk consistently in English about the semantics of English. Consequently, a great deal of effort has gone into investigating how we might restrict schema \((T)\) in some philosophically principled way, so that we could retain what is most useful and valuable about our naive understanding of truth without falling prey to antinomies.

It is interesting to inquire what happens if we throw philosophical principles to the wind and let our search for a theory of truth be guided solely by the aim of retaining as many instances of \((T)\) as we consistently can. If all we ask of our theory of truth is that it gives us as many instances of schema \((T)\) as possible, what will our theory of truth look like? \((T)\), although it is inconsistent, is such a central component of our concept of truth that it is worthwhile to try to understand how deep the inconsistency lies by inquiring how large a subset of the set of instances of \((T)\) it is possible consistently to maintain.
One reason one might undertake such an inquiry would be adherence to what Paul Horwich calls the minimalist conception of truth, according to which everything that is philosophically legitimate in our ordinary conception of truth is already implicit within schema (T). But schema (T), being inconsistent, goes well beyond the bounds of philosophical legitimacy. On a minimalist conception, all one could hope for from a legitimized theory of truth would be that it maintain as much of (T) as possible.

If one rejects minimalism, one will not suppose that (T) exhausts the naive concept of truth that we have before we confront the liar paradox, but one will nonetheless acknowledge (T) as a central component of our naive concept of truth. One will not regard it as the sole desideratum of a response to the liar paradox that it retain as many instances of (T) as possible, but one will regard it as an important constraint on a response to the liar paradox that it not restrict (T) more severely than necessary. Indeed, this is such an important constraint that it is worthwhile to study its effects by asking what our response to the liar paradox would look like if it were developed under this constraint alone, without any other considerations.

To see what a theory of truth would look like if it were developed under the sole constraint that it restrict schema (T) as little as possible, we look at maximal sets of instances of (T) that are consistent with the basic laws of syntax. Via Gödel coding, syntactic facts are intertranslatable with number-theoretic facts, so we may reformulate our task as looking for maximal sets of instances of (T) that are consistent with the basic laws of arithmetic. We take the basic laws of arithmetic to be given by some consistent arithmetical theory S which entails the axioms of Robinson's R.6

We shall find that it is easy to get7 maximal sets of instances of (T) that are consistent with S; a simple application of Zorn's lemma does the job. We shall find, in fact, that there are a great many, mutually incompatible maximal S-consistent8 sets of instances of (T). There are some maximal S-consistent sets of instances of (T) which S-entail "Tr(\(\neg 2 + 2 = 4\))," others which S-entail "Tr(\(\neg 2 + 2 = 5\))." There are some maximal S-consistent sets of instances of (T) which S-entail all instances of the schema

\[
Tr(\neg \phi \land \psi) \leftrightarrow (Tr(\neg \phi) \land Tr(\neg \psi))
\]