**Convention T and Basic Law V**

CHARLES SAYWARD

Convention T and Basic Law V of Frege’s *Grundgesetze* share three striking similarities. First, both are universal generalizations which are intuitively plausible because they have so many obvious instances. Second, both are false because they yield contradictions. Third, neither give rise to a paradox.

Suppose a person asserts that ‘The set of Fs = the set of Gs’ is to hold just in case something is an F if and only if it is a G (this is the content of Basic Law V of Frege’s *Grundgesetze*). Russell thought he had derived a contradiction from Frege’s system. This contradiction involved the notion of a set. Within that system there is the predicate ‘set not a member of itself’ and so, by Frege’s fifth axiom, a set the members of which are exactly the sets which are not members of themselves. If this set is not a member of itself...
it is a member of itself. So it is a member of itself. But if it is a member of itself then it is not a member of itself. So it is not a member of itself. So it both is and is not a member of itself.

Well, that’s a contradiction, and fairly derived from what Frege laid down (as Frege was the first to acknowledge, and freely so). There is nothing paradoxical about the fact that something of the form ‘\( \phi \) and not-\( \phi \)’ follows from something. Otherwise we should have the rain paradox: Rain is wet. It is not the case that rain is wet. Therefore, rain is wet and it is not the case that rain is wet. What is clear here is that Basic Law V appeared true to a very smart man; but, alas, it is provably false. Period. Now I want to argue that the same situation holds with respect to Convention T. While Convention T has appeared to be true to a lot of smart people, it is provably false. There is no paradox here. This is again just the case of something that sounds good (because it has a lot of plausible instances) but is, in fact, provably false.

Consider the following string of symbols:

It is not the case that the string of symbols first displayed in ‘Convention T and Basic Law V’ is true in English.

And consider the following argument: (\( \alpha \)) The string of symbols first displayed in ‘Convention T and Basic Law V’ = ‘It is not the case that the string of symbols first displayed in “Convention T and Basic Law V” is true in English’. (\( \beta \)) ‘It is not the case that the string of symbols first displayed in “Convention T and Basic Law V” is true in English’ is true in English if and only if it is not the case that the string of symbols first displayed in ‘Convention T and Basic Law V’ is true in English. Thus, (\( \gamma \)) the string of symbols first displayed in ‘Convention T and Basic Law V’ is true in English if and only if it is not the case that the string of symbols first displayed in ‘Convention T and Basic Law V’ is true in English.

Tarski reasons as follows about the argument from (\( \alpha \)) and (\( \beta \)) to (\( \gamma \)):

If we analyse this antimony in the above formulation we reach the conviction that no consistent language can exist for which the usual laws of logic hold and which at the same time satisfies the following conditions: (I) for any sentence that occurs in this language a definite name of this sentence also belongs to this language; (II) every expression formed from [‘\( x \) is true if and only if \( p \)’] by replacing ‘\( p \)’ by any sentence of the language and the symbol ‘\( x \)’ by a name of the sentence is to be regarded as a true sentence of this language; (III) in the language in question an empirically established premiss having the same meaning as (\( \alpha \)) can be formulated and accepted as a true sentence. (1956: 165)

There is nothing paradoxical about the fact that something of the form ‘\( \phi \) if and only if it is not the case that \( \phi \)’ follows from (\( \alpha \)) and (\( \beta \)). Other-
wise, we would have another rain paradox: Rain is wet. It is not the case that rain is wet. Therefore, rain is wet if and only if it is not the case that rain is wet. Tarski remarks in ‘The semantic conception of truth’: ‘It is a fact that we are here in the presence of an absurdity, that we have been compelled to assert a false sentence (since [(γ)], as an equivalence between two contradictory sentences, is necessarily false).’ (1944: 348) The key word here is the word ‘compelled’. What compels us to assert (γ)? It must be that the inference from (α) and (β) to (γ) is valid and that the two premisses are true. For an invalid argument would not compel us to assert anything. Nor would an argument with false premisses (at least one false premiss). But it is plain as anything can be that a valid argument with all true premisses has a true conclusion. It is also plain as anything can be that nothing of the form ‘φ if and only if it is not the case that φ’ is true.

Tarski reasoned this way, applying his reasoning to English: Tarski thought that English, being a colloquial language, has the feature of universality. That is, anything sayable at all is sayable in English (1956: 164). It is this feature of English which led him to conclude that (I)–(III) hold of English. Now if (I)–(III) hold of English, (α) and (β) are true. If (α) and (β) are true, so is (γ). But (γ) is a contradiction in English. Therefore, English includes a contradiction among its truths. Therefore, English is inconsistent.

The conclusion that English includes a contradiction among its truths is just a big muddle. A contradiction is false. Nothing is both true and false. Therefore, no language includes a contradiction among its truths. Nothing could be simpler.

Where Tarski went wrong is in making this tacit inference: English is a colloquial language. Every colloquial language has the feature of universality. Therefore, (I)–(III) hold of English. The conclusion does not follow. (I) and (III) hold of English, but (II) does not hold.

Here is how Tarski should have reasoned: The inference from (α) and (β) to (γ) is valid. It is easy to empirically determine that (α) is true. It is a matter of logic that (γ) is not true. Therefore (β) is not true. (β) is an instance of (II), which is Convention T. Therefore, Convention T is not true. This seems to me to be about as convincing as anything can be in philosophy.

There is no paradox here. All we have is another case of a very smart man (in this case Tarski) thinking that some universal generalization (in this case Convention T) is true when in fact it is provably false.
References
