HOW NOT TO STATE THE T-SENTENCES

Volker Halbach

1

On many accounts of truth the T-sentences or similar equivalences feature prominently.[1] The T-sentences are mostly seen as sound principles, although there is a disagreement about whether the T-sentences capture the notion of truth completely.

There is an obvious obstacle in the formulation of these equivalences: certain instances of them lead to inconsistencies. There have been various proposals to overcome this obstacle by restricting the class of legitimate instances. For instance, one might feel tempted to allow only those instances of

(T) $$\mathsf{T}\ulcorner A \urcorner \equiv A$$

that do not allow one to derive inconsistencies. McGee (1992) has shown that this strategy is doomed for various reasons: there are many such maximal inconsistent sets of instances and they may have absurd consequences that are as bad as an outright inconsistency. There are more proposals for restrictions on A in (T) that have their own problems. However, almost all authors agree that there is a fall-back position that might be too restrictive but that saves (T) from inconsistency: one bans all A containing the truth predicate from (T).

In the following I shall put forward an argument showing that even this restrictive and putatively safe way to state the T-sentences may lead to paradox and may require further provisos, at least in certain situations.

---

[1] I am using the term 'T-sentence' in a wide sense. I do not claim that this use tallies with Tarski's use of the term.

Besides the truth predicate we shall want to have other predicates for necessity, analyticity and the like, because we surely want to be able to say 'There are a priori truths that are not necessary' or 'All analytic truths are necessary.'

Montague (1963) has shown that modal predicates have also their problems. In a nutshell, he showed that an inconsistency still arises if the right-to-left direction of (T) is weakened to a mere rule and the other direction is retained as an axiom scheme. More precisely, on the background of a theory allowing for diagonalisation the scheme $\mathsf{N}\ulcorner A\urcorner \supset A$ is inconsistent with the rule of necessitation that allows one to deduce $\mathsf{N}\ulcorner A\urcorner$ if A has been derived.

Montague's observation has been taken by many authors, including Montague himself, as an argument against the predicate conception of modalities and in favour of the operator approach. Here I do not want to go into the discussion about the operator conception versus the predicate conception, but on the operator approach it is difficult to formalise quantified statements like 'There are a priori truths that are not necessary' or 'All analytic truths are necessary.' This suggests employing the same strategy as in the case of the truth predicate: the instances of the axiom scheme $\mathsf{N}\ulcorner A\urcorner \supset A$ and the rule of necessitation are restricted to As that do not contain the predicate N.

In the case of N we should surely like to have axioms in addition to the scheme and the rule I have just expounded that allow us to characterise N more specifically as necessity, analyticity or whatsoever. But these additional axioms are not needed in the following, so we focus on the very rudimentary theory given by $\mathsf{N}\ulcorner A\urcorner \supset A$ and necessitation.

This theory looks sound: $\mathsf{N}\ulcorner A\urcorner \supset A$ is a highly plausible axiom scheme for necessity and other modalities. If the theory does not contain any contingent axioms, necessitation is also hardly disputable. And since in the axiom scheme and the rule of necessitation only sentences not containing N are allowed, this theory should be as safe from inconsistency as the analogously restricted T-sentences.

3

Let Σ be the theory that is given by the axioms of a theory like Robinson's arithmetic Q that allows for a proof of the diagonal lemma, and that

comprises the following axioms and the following rule:

(T) $\mathsf{T}\ulcorner A\urcorner \equiv A$, where A does not contain T

(N1) $\mathsf{N}\ulcorner A\urcorner \supset A$, where A does not contain N

(N2) From A you may infer $\mathsf{N}\ulcorner A\urcorner$, if A does not contain N

Then Σ is inconsistent. To see this, reason in Σ as follows:

$$
\begin{array}{ll}
D \equiv \sim\mathsf{T}\ulcorner\mathsf{N}\ulcorner D\urcorner\urcorner & \text{diagonalisation} \\
\mathsf{T}\ulcorner\mathsf{N}\ulcorner D\urcorner\urcorner \equiv \sim D & \\
\mathsf{N}\ulcorner D\urcorner \supset \sim D & \text{(T)} \\
\mathsf{N}\ulcorner D\urcorner \supset D & \text{(N1)} \\
\sim\mathsf{N}\ulcorner D\urcorner & \text{two previous lines} \\
\sim\mathsf{T}\ulcorner\mathsf{N}\ulcorner D\urcorner\urcorner & \text{(T)} \\
D & \text{first and last line} \\
\mathsf{N}\ulcorner D\urcorner & \text{(N2)}
\end{array}
$$

The last line contradicts the fourth line from the bottom.

Of course, each of (N1)+(N2) and (T) taken separately are consistent with Q, respectively, but joining them into one single theory spells disaster.

The inconsistency does not contradict the Joint Consistency theorem (or the Interpolation theorems), because the theory (T) for truth contains axioms with occurrences of N and (N1)+(N2) contain axioms with occurrences of T, although (T) is decidedly a truth-theoretic axiom scheme and not a scheme concerning the modality N.[2] If (T) were also conceived as an axiom scheme for N, one could equally claim that the T-sentences are axioms concerning *all* predicates that may occur in A. And it would be strange to call (T) a theory of snow, because the sentence A may contain this term.

4

In order to restore consistency, one might want to say that an inappropriate notion of occurrence has been employed in the restriction of instances of (N1)+(N2) and (T). It could be argued that although the sentence $\mathsf{N}\ulcorner D\urcorner$ in the proof above does not contain an occurrence of T in the usual sense, it

---

[2]See, e.g., (Chang and Keisler 1990) for a statement and proof of the Joint Consistency theorem.

contains an occurrence of T is another relevant sense: T occurs in $\mathsf{N}\ulcorner D\urcorner$ indirectly, because T occurs in D. Therefore, or so one might argue, $\mathsf{N}\ulcorner D\urcorner$ should not be admitted as an instance of (T). Thereby one might hope to restore the consistency of Σ by postulating that in (T) the set of possible instances is restricted to those A that contain T neither directly or indirectly.

In order to exclude instances of (T) where A contains T indirectly, one will have to disallow also singular terms in A that refer to sentences containing T; moreover one will have to rule out instances of (T) where A embraces quantification over sentences with T.

If the language contains, for instance, definite descriptions whose reference depends on contingent factors, the set of legitimate instances of (T) depends on contingent factors, thereby rendering the whole theory of truth variable depending on contingent matters. The restriction of quantifiers will add further problems. Even in the absence of singular terms with contingent reference it will hardly be possible to provide a recursive set of axioms. For determining whether a quantification ranges over sentences containing the truth predicate may be difficult to determine. Therefore working out the details of this strategy to rule out instantiating formulae A with indirect occurrences of T is messy, and it is hard to see how it should yield a sensible set of axioms for truth. Here Leitgeb's (2005) notion of dependence might be useful, but it will hardly resolve the problems addressed here.

The failure of this approach suggests a different type of restriction on (T): A is disallowed to contain any occurrences of T *or* N.[3] Thereby the third line in the above derivation of the inconsistency would be blocked: $\mathsf{T}\ulcorner\mathsf{N}\ulcorner D\urcorner\urcorner \equiv \mathsf{N}\ulcorner D\urcorner$ would cease to be an legitimate instance of (T), because it contains an occurrence of N.

This approach would block virtually any interaction of truth and necessity and probably many other notions. Moreover, the T-sentences would have to be drastically further restricted: not only would T be banned from the set of legitimate instances but also many further predicates. It is, however, hard to tell which predicates should be put on the Index. The inconsistency I have addressed might be only one among many.[4]

I find it therefore hard to see how the schema (T) could sensibly be for-

---

[3] Analogous points could be made for (N1)+(N2), and the consistency could be restored by restricting the possible instances of (N1)+(N2) to those where A contains nether T nor N. In this case the following remarks would apply *mutatis mutandis*.

[4] Not much is known about inconsistencies arising from the interaction of two 'modal' predicates. For exceptions see (Niebergall 1991), (Halbach 2002), (Horsten 2002), (Horsten and Leitgeb 2001).

mulated without imposing very trenchant restrictions on possible instances. These restrictions must be much more discriminating than the usual proviso 'and *A* must not contain the truth predicate.'[5]

References

Chang, C. C. and H. J. Keisler. 1990. *Model Theory*. 3rd. North Holland.

Halbach, V. 2002. Modalized Disquotationalism. In *Principles of Truth*, edited by Volker Halbach and Leon Horsten, 75–101. Frankfurt a.M.: Dr. Hänsel-Hohenhaus.

Horsten, L. 2002. An Axiomatic Investigation of Provability as a Primitive Predicate. In *Principles of Truth*, edited by Volker Halbach and Leon Horsten, 203–20. Frankfurt a.M.: Dr. Hänsel-Hohenhausen.

Horsten, L. and H. Leitgeb. 2001. No Future. *Journal of Philosophical Logic* 30:259–65.

Leitgeb, H. 2005. What Truth Depends On. *Journal of Philosophical Logic* 34:155–92.

McGee, V. 1992. Maximal consistent sets of instances of Tarski's schema (T). *Journal of Philosophical Logic* 21:235–41.

Montague, R. 1963. Syntactical Treatments of Modality, with Corollaries on Reflexion Principles and Finite Axiomatizability. *Acta Philosophica Fennica* 16:153–67. Reprinted in (Montague 1974, 286–302).

———. 1974. *Formal Philosophy: Selected Papers of Richard Montague*. New Haven and London: Yale University Press. Edited and with an introduction by R. H. Thomason.

Niebergall, K.-G. 1991. Simultane objektsprachliche Axiomatisierung von Notwendigkeits- und Beweisbarkeitsprädikaten. Master's thesis, Ludwigs-Maximilians-Universität München.

New College
Oxford OX1 3BN, *United Kingdom*
volker.halbach@philosophy.ox.ac.uk