Philosophical Communications, Web Series, No. 35, pp. 203-230
Dept. of Philosophy, Göteborg University, Sweden
ISSN 1652-0459

Kvantifikator för en Dag

Essays dedicated to Dag Westerståhl on his sixtieth birthday

‘Means that’ and the liar*

Peter Pagin

Abstract

By means of ‘means that’ and propositional quantification, we can define a truth predicate. This also allows the construction of liar sentences, either by self-reference or by means of quantification. In order to avoid inconsistency, restrictions on expressive power must be imposed, and the question is how far such restrictions will limit our ability to say of what is intuitively described as “meaningful” that it is precisely meaningful.

1. ‘Means that’ and truth predicates

We like to think of the sentences we use as meaningful. We like to think that for any well-formed sentence of our natural language, be it Swedish, English, Tagalog or any other of the terrestrial languages, a speaker can utter that sentence and in virtue of semantic properties already established, thereby express a proposition, something that is true or false depending on what the world is like. We would not like to be in the predicament of constructing a sentence that is well-formed, as far as we can tell, and for which we can even give argument that should lead us to accept it as intuitively correct, and yet at the same time have to judge it meaningless, in the sense of not expressing any proposition. Yet, there seem to be such sentences, or at least nearly enough.

*An earlier version of this paper was presented at the LLS Colloquium at Ohio State University. I had extremely valuable comments from Neil Tennant, Kevin Sharp, Gabriel Uzquiano Cruz, Craige Roberts, and -- especially -- Stewart Shapiro. Before that I received helpful comments from Sten Lindstöm on a first draft. I'd also like to take the opportunity to thank Dag for uncountably many stimulating discussions over the years, for great cooperation and great friendship.

The basic fact leading to this problem is that the vocabulary that we use to express our judgments about meaningfulness is itself part of the vocabulary we that we pronounce on. That a situation of this kind can give rise to problems is well known from the history of the semantic paradoxes, and in particular the liar paradox itself, where part of the semantic vocabulary. 'true', is used to generate a contradiction. We have a basic liar sentence in

(1) (1) is not true

To get a contradiction out of (1) we also need a disquotation principle:

(DQ) $\ulcorner q \urcorner$ is true iff q

i.e. a schema with instances such as

(2) 'snow is white' is true iff snow is white

Then, since we have the identity

(3) $(1) =$ '(1) is not true'

by means of (D) and classical logic, we get the familiar result that if (1) is true iff (1) is not true, and hence that it is both true and not true. Even though (D) is suspect because of the contradiction, it has high independent plausibility, and therefore giving it up is not an easy solution.

The liar paradox soon becomes relevant when thinking about meaningfulness, since with the help of the relational expression 'means', or 'means that', and propositional quantification, we can define a truth predicate for sentences.[1] In natural English we could state it as

$$\text{For some proposition } p, \ldots \text{ means that } p, \text{ and it holds that } p$$

[1] I am here going to focus on 'means that', which can be described as an operator which, when applied to a sentence forms a one-place predicate. 'Means', on the other hand, is more naturally described as a two-place predicate, where the second argument place gets filled by singular terms referring to propositions, and then *that*-clauses are treated precisely as such singular terms. Since in the present context, I can make do with substitutional quantification, which is less controversial than objectual propositional quantification, it is natural to prefer the 'means that' alternative.

That this is a truth predicate is intuitively clear from examples such as

(4) 'Snow is white' means that snow is white, and snow is white

(5) For some proposition p, 'Snow is white' means that p,

Here it is clear that the sentence 'Snow is white' is true, by normal standards, if (4) is true. Moreover, whatever proposition 'Snow is white' can correctly be said to mean (in English), it is such that the sentence is again intuitively true if that proposition holds.[2] So, (5) can be be treated as a truth predicate: a sentence is true just in case it means something that holds. If we allow ourselves formal propositional quantifiers and sentential logical constants, we can do without the 'holds that', and simply have

$$\exists p(\ldots \text{ means that } p \ \& \ p)$$

I shall here treat the quantification as substitutional. That is, a sentence $\exists pA$ is true just if there is a true substitution instance $A(q/p)$, where all (free) occurrences of 'p' in A are replaced by 'q'.

I do not claim that this is the correct, or the best, or even, all things considered, *a* correct explication of the meaning of 'true' and other natural language (sentential) truth predicates. It is *a* truth predicate in the sense of satisfying to a sufficient degree the pattern of application of truth predicate. Since this is a truth predicate, we can form a liar sentence with it:

(6) $\neg\exists p((6) \text{ means that } p \ \& \ p)$

We could avoid this particular liar sentence by e.g. ruling out the kind of self-reference exemplified, but on the one hand this does not help against liar sentences where the liar sentence is identified by means of a unique property ('there is a unique sentence that is Φ and ...'), and on the other hand, ruling out a particular linguistic device entirely needs a much better reason than that of avoiding a bad consequence of using it, together with other devices, in some particular applications. So in general, attempts to avoid semantic paradoxes

[2] I have used 'it holds that' rather than 'it is true that' only in order to reserve 'true' for sentences in this context.

in natural language by means of linguistic legislation, is something to be avoided as far as possible. This holds both of self-reference and of propositional quantification.

If these advices are accepted, we are free to form liar sentences by means of 'means that', and as a consequence there will be certain limitations in its applicability. In the remainder of this paper I shall try to determine what those limitation are, and whether we can live with them.

2. Incomplete internal semantics

The use of 'means that' for defining a truth predicate is not new. It is explored in Parsons 1974, in the format

(CP) T(y) iff $\exists x$(x is a proposition $\&$ y expresses x $\&$ x is true)

(Parsons 1974, 389), for defining a sentential truth predicate from a propositional truth predicate.[3]

Something very similar is also used in Glanzberg 2004:

(G) $\exists p(\mathrm{Expresses}(s,p,c)\ \wedge\ \mathrm{True}(p))$

(Glanzberg 2004, 32).[4]

Both Parsons (1974, 387-89) and Glanzberg (2004, 33-34) use their definition of truth to perform liar reasoning, very close to what follows. In so doing they make use of a disquotation principle (Parsons 1974, 387; Glanzberg 2004, 33), equivalent to

(D) $\forall p$(if $\ulcorner q \urcorner$ means that p, then p iff q)

[3] I have changed the notation somewhat. It should be noted that in his footnote 12, Parsons also considers a formulation without a propositional truth predicate, formulated in Montague's IL.

[4] A related but non-sentential alternative can be found in Hill 2002, 22:

(CH) For any object x, x is true iff $\Sigma p(x = p\ \&\ p)$

where 'Σ' is a substitutional quantifier. Except for the detail that (T) does not contain the variable 'p' in what appears to be a singular term position, the differences are that (T) defines a truth predicate of sentences, while the truth predicate of (CH) only has propositions in its extension, and that the quantifier of (T) is objectual, with the variable 'p' ranging over propositions.

This is a *conditional* principle of disquotation, allowing us to infer that q from the premise that $\ulcorner q \urcorner$ is true, but not that $\ulcorner q \urcorner$ is true from the premise that q, unless we also have the premise that $\ulcorner q \urcorner$ means something. As will be seen shortly, this is crucial.

I shall here assume that the usual generalization and instantiation laws are valid for the propositional quantifiers '$\exists p$' and '$\forall p$'. Now consider the sentence

(7) $\neg\exists p((7) \text{ means that } p \;\&\; p)$

We shall perform a liar reasoning with (7). We start by assuming

(8) $\exists p((7) \text{ means that } p)$

We have as an instance of (D)

(9) $\forall q($ if '$\neg\exists p((7) \text{ means that } p \;\&\; p)$' means that q, then
q iff $\neg\exists p((7) \text{ means that } p \;\&\; p))$

We have the identity

(10) $(6) =$ '$\neg\exists p((7) \text{ means that} p \;\&\; p)$'

and hence

(11) $\forall q($ if (7) means that q,
then q iff $\neg\exists p((7) \text{ means that } p \;\&\; p))$

Now assume

(7) $\neg\exists p((7) \text{ means that } p \;\&\; p)$

From (7) and (11) we may infer

(12) $\forall q(\text{if } (7) \text{ means that } q, \text{ then } q)$

And from (8) and (12) we can infer

(13) $\exists p((7) \text{ means that } p \,\&\, p)$

From (7) and (13) we can infer

(14) $\exists p((7) \text{ means that } p \,\&\, p)$

discharging the assumption (7). But from (14) and (11) we get

(15) $\forall p(\text{if } (7) \text{ means that } p, \text{ then } \neg p)$

(15) contradicts (14). This allows us to conclude

(16) $\neg\exists p((7) \text{ means that } p)$

discharging the assumption (8). We have proved that there is no proposition that (6) expresses. This is a desirable result, which thanks to the restriction of the disquotation principle avoids contradiction.

However, from (16) we may again infer

(17) $\neg\exists p((7) \text{ means that } p \,\&\, p)$

(17) is identical with (7). We have thus proved (7)[5]. So far we don't have a contradiction (note that we need (8) for deriving (13)). But it is natural to assume the following principle:

(P) For any sentence s, if s is provable, then $\exists p(s \text{ means that } p)$.

We add our reflection over (7) and (17):

(18) (7) is provable

and from (18) and the principle (P) we infer

[5]This step is explicitly taken in Glanzberg 2004, 34.

(19) $\exists p((7) \text{ means that } p)$

which contradicts (16). So we have an inconsistency.

Beside the logic and the self-reference, the contributors to the contradiction are the two principles (D) and (P). Since there is no reason to call the logic into doubt, and self-reference is acceptable, and since (D) can hardly be rejected, we must focus on (P). In an intuitive sense (P) is valid, for (7) must be meaningful, i.e. express a proposition, given that it is provable. On the other hand, if (P) is accepted straight off, together with the rest of our ingredients, we have an inconsistency.

A well-established way of resolving such difficulties the tradition of modern logic is to distinguish between an internal and an external perspective. This is typically applied in set theory. For instance, we need to distinguish between the claim that there is a set of all sets that are not members of themselves, as expressed *in* a particular set theory about the sets in its domain, and the corresponding claim expressed in a *meta-language*. The crucial difference between the two existence claims is that the meta-linguistic claim is made with respect to a wider domain of quantification. If we make the domain explicit, and let $U \subset U'$, we can represent the internal and the external claim, respectively, as:

(20) a. $\exists x \in U(\forall y \in U(y \in x \leftrightarrow \neg(x \in x)))$

b. $\exists x \in U'(\forall y \in U(y \in x \leftrightarrow \neg(x \in x)))$

Here (20a) leads to Russell's paradox but (20b) does not, since the Russell collection x need not there be in U itself. In case the domain of quantification is *not* made explicit, the two statements will look syntactically the same, but their content will be different because of an implicit contextual element, the domain of quantification, with respect to which they differ. We would then say that there is a difference in *context* that resolves the issue: in a context with a wider domain of quantification the existence claim can be consistently made.

Both Parsons and Glanzberg appeal to such a context shift for resolving the problem above. We seem at the same time to be forced to conclude that a particular sentence does not express any proposition, and yet this sentence seems perfectly well-formed and even assertible, and hence should be deemed after all to express a proposition. The contradiction can be resolved by a contextual shift from a smaller to a wider domain of propositions. When we first conclude with (16) that there is no proposition that (7) expresses, this may be taken to be correct with respect to the domain of propositions over which the variable 'p' ranges, and when we conclude with (19) that (7) does express a proposition after all, we have a larger domain in mind. In the larger domain, there is a proposition expressed by (7), whereas in the smaller there isn't. Between (16) and (19) there has been a tacit contextual shift from a smaller to a larger implicit domain of quantification, and hence there is no contradiction. This is in essence what Parsons and Glanzberg propose.[6]

Accounts of this kind have obvious advantages. They provide not only a solution to the problem of some particular liar sentence but offer a strategy with dealing with new ones that will arise. For with the increased domain a new and stronger liar sentence can be produced, and it will be handled by means of yet another domain extension.

One serious problem with the approach is that it relies on the assumption that the initial domain is limited and can be expanded. If the speaker does quantify of *everything* to begin with, then the solution is not available. Glanzberg does suggest that only a limited domain of propositions is available for the speaker, and that it is limited by the resources the speaker has for expressing propositions. With the expression of a new liar sentence a new semantic relation gets added to the context, and then new propositions can be constructed, and the domain increases (Glanzberg 2004, 41-44).

However, even if this does account for some liar sentence utterances, it would seem that we are not always so limited as speakers. In the context of setting the theory we seem to be quantifying over a domain of domains of proposition, say U, with element domains $u$and if we can do that we can exploit the feature for a strengthened liar:

(21) $\neg\exists u \in U(\exists p \in u((21) \text{ means that } p \,\&\, p))$

As is often remarked in the liar literature, the formulation of general theory provides the resources for constructing a "superliar" sentence that cannot be acceptably handled by the

[6]Parsons's work was a pioneering one and Glanzberg's much later work appeals to a lot of later linguistic literature for motivating what he calls 'extra-ordinary context dependence', and employs a sophisticated model theory based on the theory of admissible sets.

theory. In case a very far-fetched construction is needed, like a sentence involving quantification over all levels in a Tarskian hierarchy of languages, we may claim to have achieved something of value, since only very special constructions are outside the scope of the account. But this does not seem to be the case for an account that builds on a restriction of the domain of propositions. The idea of *all propositions* does seem to be a simple one, and there does not seem to be any good reason for claiming that I cannot quantify over propositions I cannot express.

Parson (1974, note 13) claims that discourse with a generality that transcends all set domains must have a certain kind of systematic ambiguity, and that we cannot say what this generality consists in except in a language having such a systematic ambiguity. Such discourse does not employ absolutely general quantification. Burge, another contextualist, rejects the possibility of quantifying over all contexts (Burge 1979, 192), while this is accepted in the contextualist account of Simmons 1993 (171-76). According to Simmons, a sentence involving quantification over all contexts marks the limit of the contextualist account.

There are indeed difficulties with the idea of a total collection of all propositions. Russell himself rejected the idea of a well-defined totality of all propositions (Russell 1908, 157). More recently, Sten Lindstr"om has suggested that the resolution of Kaplan's problem of plenitude (Kaplan 1994, Lindstrom 2003a), as well as of the Russell-Myhill antinomy (Lindstrom 2003b), consists in the rejection of the idea that we can quantify over all propositions.[7]

On the other hand, there severe difficulties in *stating* the view that quantification must be restricted, for a sentence like

(22) I do not quantify over everything

[7]Both paradox derive from defining properties for propositions. In Kaplan's problem of plenitude there is a property Q such that for each proposition p there is a possible world w where p is the unique proposition having Q: because of this there cannot be more propositions than worlds and on the other hand, since propositions are sets of worlds, by Cantor's diagonal proof there must be more sets of worlds than worlds.

In the Russell-Myhill antinomy, there is an analogous cardinality problem involved with the idea that propositions are structured. Then, on the one hand there must be at least as many propositions as there are properties, for if Φ and Ψ are different properties, $\Pi\Phi$ and $\Pi\Psi$ must be different propositions, where Π means 'everything has'. On the other hand, by the diagonal argument for properties, there are more properties of things than there are things, including propositions.

is false, since 'everything' comprises exactly everything I quantify over. This problem has been recently pressed by Tim Williamson (2003) on behalf of the generality absolutist view that quantification is unrestricted. That is, according to the generality absolutist view, we do quantify over everything.[8]

Because of these difficulties, it is worth while trying another alternative, more following Tarski in respect of focusing on restrictions of the application of predicates, in this case 'means that'.[9]

Given an informal notion of *meaning something*, or *expressing a proposition*, (P) must be considered intuitively correct, and this is a judgment passed from the external perspective. On the other hand, from the fact that a sentence s i provable we cannot legitimately infer that the two-place predicate '. . . means that . . . ' that is actually in use is applicable to s. The predicate actually in use, in this case, is the predicate occurring in (7) itself. We therefore distinguish between a predicate 'means that_0' of a somewhat regimented language L, and an informal unregimented predicate 'means that', without subscript, that belongs to a meta-language ML. We identify 'means that_0' with 'means that' as occurring in (7)--(19).

The upshot is then that the *internal semantics* of L, expressed within L itself, is incomplete. There is a *means that* relation holding between (7) and the proposition *that* $\neg\exists p((7) \text{ means that}_0\ p\ \&\ p)$, but that relation is only partially expressed by 'means that_0'. The pair

$$\langle (7),\ \text{the proposition that } \neg\exists p((7) \text{ means that}_0\ p\ \&\ p)\rangle$$

is not in the extension of 'means that_0'.

[8]Williamson gave this paper in the *Logic and Language* conference in Birmingham 2003, and I had been asked to comment. I claimed that there are corresponding difficulties of thesis formulation for the absolutist. The sentence

(23) We quantify over everything

is bound to be true, whether the domain of quantification is unrestricted or not. So there is a problem of stating the view both for the generality relativist (or restricitvist) and for the generality absolutist. On the other hand, it is better that one's attempt to state one's position results in a true claim than in a false one.

[9]Note that the initial question of expressing intuitive verdicts about meaningfulness would still stand if instead we did pursue the alternative of restricting quantification domain.

I shall now sketch a semantics for a language L_0 that is an impoverished formal counterpart to L.

3. Modeling incomplete internal semantics

The basic idea for a semantics with an incomplete internal 'means that' predicate is to impose a non-circularity condition in the truth definition clause for the formal counterpart of 'means that'. This is done by means of a definition of *dependence*, below. Propositional quantification will be substitutional. We consider a language L_0:

Definition 3.1. The language L_0 consists of

a. denumerably many sentence letters 's_1', 's_2', ...
b. denumerably many primitive singular terms '(1)', '(2)', ...
c. denumerably many propositional variables 'p_1', 'p_2', ...
d. propositional logic constants '$\neg$' and '&'
e. quantifier symbol '$\exists$'
f. the relational expression 'R'
g. quote sign '$\bar{\ }$'

Since we can form new formulas by means of terms, and new terms by means of sentences, we need to define formulas, terms and sentences together. Below I shall rely on the ordinary conception of free and bound variables.

Definition 3.2. The set of well-formed singular terms (wfts) and the set of well-formed formulas (wffs) of L_0 are simultaneously inductively defined as the smallest sets satisfying:

a. every sentence letter of L_0 is a wff
b. every propositional variable of L_0 is a wff
c. if ϕ is a wff, then $\ulcorner \neg\phi \urcorner$ is a wff
d. if ϕ and ψ are wffs, then $\ulcorner \phi \ \& \ \psi \urcorner$ is a wff
e. if ϕ is a wff and q a propositional variable, then $\ulcorner \exists q\phi \urcorner$ is a wff
f. a wff ϕ is a sentence iff ϕ does not contain any free variables
g. if ϕ is a sentence, then $\ulcorner \bar{\phi} \urcorner$ is a singular term

h. if t is a singular term and ϕ a formula, then $\ulcorner tR\phi \urcorner$ is a wff

In order to render the internal *means that* predicate, R, incomplete, we make sure that terms and sentences that involve circularity are excluded. For this purpose we define the notion of being *self-dependent*:

Definition 3.3. Dependence:

a. A sentence s depends on a sentence s' if there is a term t that occurs in s and t denotes s'

b. If a sentence s depends on a sentence s' and s' depends on a sentence s'', then s depends on s''

c. A sentence is *self-dependent* iff it depends on itself by clauses (a)–(b)

Definition 3.4. A model $\mathcal{M}$ for L_0 is a pair $\langle I, V\rangle$, where I is an interpretation function that provides denotations to simple terms and truth values (T, F) to sentence letters and V is such that

a. For a primitive term t, $V(t) = I(t)$

b. $V(\bar{s}) = s$

c. For a sentence letter $s, \mathcal{M} \models s$ iff $I(s) = \mathrm{T}$

d. $\mathcal{M} \models s\&s'$ iff $\mathcal{M} \models s$ and $\mathcal{M} \models s'$

e. $\mathcal{M} \models \neg s$ iff not $\mathcal{M} \models s$

f. $\mathcal{M} \models tRs$ iff $V(t) = s$ and s is not self-dependent.

g. $\mathcal{M} \models \exists p\phi$ iff for some substitution instance $\phi(s/p)$, $\mathcal{M} \models \phi(s/p)$

By means of clause f, in a sense we let the meaning of a sentence be itself, since we don't allow any truth preserving substitution (other than the identity substitution) on the right hand side of R. Hence, there will be no non-trivial synonyms. This simplifies things enormously (both in the good and in the bad sense).

We can verify from this truth definition that the expression `$\exists p(\ldots Rp \ \& \ p)$' (and the variants we get by substitution of bound variables) behaves like a truth predicate. For if

(24) $\mathcal{M} \models \exists p(tRp \,\&\, p)$

then by clause g of 3.4, there is a sentence s in L_0 such that

(25) $\mathcal{M} \models tRs \,\&\, s$

and therefore a sentence s such that $V(t) = s$ and $\mathcal{M} \models s$. That is, what t denotes is true in $\mathcal{M}$.

Conversely, suppose that $\mathcal{M} \models s$ and that for some term t, $\mathcal{M} \models tRs$. We then get (24) by conjunction introduction and existential generalization.

Similarly, we can see that the regimented counterpart to the disquotational principle (D), i.e. the schema

(D+) $\neg\exists p(\bar{s}Rp \,\&\, \neg(\neg(p\&\neg s) \,\&\, \neg(s\&\neg p)))$

is valid. For assume that this is false. Then there are sentences s and s' such that

(26) $\mathcal{M} \models \bar{s}Rs' \,\&\, \neg(\neg(s'\&\neg s) \,\&\, \neg(s\&\neg s'))$

But then $\mathcal{M} \models \bar{s}Rs'$, and hence $V(\bar{s}) = s'$. Since $V(\bar{s}) = s$, we can infer that $s = s'$. But then the second conjunct of (26) is false, i.e.

(27) not $\mathcal{M} \models \neg(\neg(s'\&\neg s) \,\&\, \neg(s\&\neg s'))$

and (26) itself is false. So the disquotational principle holds.

We can evaluate the liar sentence

(28) $\neg\exists p((28)Rp \,\&\, p)$

as follows:

(29) $\mathcal{M} \models \neg\exists p((28)Rp \,\&\, p)$

iff there is no sentence s in L_0 such that

(30) $\mathcal{M} \models (28)Rs \ \&\ s$

But for

(31) $\mathcal{M} \models (28)Rs$

to be true it must hold that and s isn't self-dependent. Since $V((28))$=‘$\neg\exists p((28)Rp \ \&\ p)$’, (28) is self-dependent. So there is no sentence s satisfying (30), and hence (29) is true, i.e. (28) is true in $\mathcal{M}$. Analogously, we can see that the truth teller

(32) $\exists p((32)Rp \ \&\ p$

is false in $\mathcal{M}$.

The intuitive reasoning by which (7) was informally proved is represented in $\mathcal{M}$, insofar as the truth of (28) depends on the falsity of (31).

L_0 does not contain its own truth predicate, since a sentence predicating truth of the liar sentence (28), like

(33) $\exists p((28)Rp \ \&\ p)$

is false in $\mathcal{M}$. Note that

(34) $\exists p((33)Rp \ \&\ p)$

is false i in $\mathcal{M}$ as well, because of the falsity of (33), not because (34) itself is self-dependent, for it isn't: (28) does not depend on (33).

The model seems to preserve both consistency and bivalence. However, since the liar sentence (28) is true in $\mathcal{M}$, and the informal liar sentence (7) is informally provable, the internal truth predicate only approximates what we want. Can we represent the truth of (28) in the object language itself?

4. Extending the internal semantics

In order to approximate more closely the intuitive truth predicate we want to accommodate in some way the reflection on truth and provability stated in

(P) For any sentence s, if s is provable, then $\exists p(s \text{ means that } p)$.

One way of doing this is accept (P) in a *diachronic* sense: If we realize at a stage δ that a sentence $\ulcorner q \urcorner$ is provable, we have the right to extend the *means that* relation current at δ, i.e. its expression in the language itself 'means that' at δ), so that at stages after δ, the pair

$$\langle \ulcorner q \urcorner, \text{ the proposition that } q \rangle$$

is included in the *means that* relation, which it wasn't before. With transitions of this kind we can extend the *means that*-relation as expressed in the language itself in various ways.

We can implement this idea by introducing a second internal *means that* predicate, that has virtually the same semantics as 'means that$_0$', except that in virtue of being different, it can represent the intuitive *means that* relation in handling sentences with occurrences of 'means that$_0$', such as the original liar sentence. For we can accept

(35) $\exists p((7) \text{ means that}_1 \text{ \& } p)$

since (35) does not contradict (7) itself. Hence, the pair

$$\langle (7), \text{the proposition that } \neg\exists p((7) \text{ means that}_0\ p \text{ \& } p)\rangle$$

is in the extension of 'means that$_1$'

Together with the new predicate we need a new disquotation principle

(D_1) $\forall p(\text{if } \ulcorner q \urcorner \text{ means that}_1 p, \text{ then } p \text{ iff } q)$

And, of course, once we have (D_1) we can form a new liar sentence by means of 'means that$_1$':

(36) $\neg\exists p((36) \text{ means that}_1\ p \text{ \& } p)$

and we can perform the same reasoning as above with our old liar sentence(7). We do not, however, immediately need a third *means that* predicate in order to express the intuitive truth of (36), for we can do this by means of our original 'means that$_0$':

(37) $\exists p((36) \text{ means that}_0\ p \ \&\ p)$

since (37) does not contradict (36). Hence, the pair

$$\langle (36), \text{the prop that } \neg\exists p((36) \text{ means that}_1\ p \ \&\ p)\rangle$$

is in the extension of 'means that$_0$', after the introduction of the new predicate. This shows that we do not automatically get a hierarchy by the introduction of a new predicate to handle the original liar. Rather, the two *means that* predicates interact in an unordered way, much as is pictured by the context anchored truth predicates in the contextual theories of Burge and Simmons.

But even if there is no semantic order between 'means that$_0$' and 'means that$_1$', we can form a superordinate predicate by means of taking their disjunction. This gives us a superordinate truth predicate

(38) $\exists p((\ldots \text{ means that}_0 \text{ or } \ldots \text{means that}_1\ p) \ \&\ p)$

Clearly, any sentence satisfying either of the simple truth predicates also satisfies the disjoined one. However, with the disjoined, strengthened truth predicate we also have new strengthened liar sentences, such as:

(39) $\neg\exists p(((39) \text{ means that}_0\ p \text{ or } (39) \text{ means that}_1\ p) \ \&\ p)$

To show that this is a strengthened liar, assume that (39) is meaningful and true (0). That is, assume

(40) $\exists p((39) \text{ means that}_0\ p)$

(41) $\exists p((39) \text{ means that}_0\ p \ \&\ p)$

We have as an instance of (D)

(42) $\forall q$(if '$\neg\exists p$((39) means that$_0$ p & p)' means that$_0 q$,

then q iff $\neg\exists p$(((39)means that$_0$ p or means that$_1$ p) & p))

We have the identity

(43) (39) = '$\neg\exists p$(((39) means that$_0$ p or (39) means that$_1$ p) & p)'

and hence

(44) $\forall q$(if (39) means that$_0 q$,

then q iff $\neg\exists p$((39) means that$_0$ p or means that$_1$ p) & p))

From (41) we can infer

(45) $\exists p$(((39) means that$_0$ p or (39) means that$_1$ p) & p)

and from (44) and (45) we have

(46) $\forall q$(if (39) means that$_0 q$, then $\neg q$)

which is to say

(47) $\neg\exists p$((39) means that$_0$ p & p)

and since (47) is derived from its own negation as assumption, we can discharge the assumption and repeat (47).

With symmetrical reasoning we can infer

(48) $\neg\exists p$((39) means that$_1$ p & p)

and from (47) and (48) together we can infer

(49) $\neg\exists p$(((39) means that$_0$ p or (39) means that$_1$ p) & p)

which is to say (39) itself. Then, by means of (44) and (39) we have

(50) $\forall q(\text{if (39) means that}_0\ q,\ \text{then}\ q)$

And from (40) and (50) we can infer

(51) $\exists p(\text{(39) means that}_0\ p\ \&\ p)$

contradicting (39). We then have to conclude that

(52) $\neg\exists p(\text{(39) means that}_0\ p)$

discharging the assumption (40). By symmetric reasoning we can conclude

(53) $\neg\exists p(\text{(39) means that}_1\ p)$

That is, (39) is intuitively meaningful and true but does not fall under either of the two internal 'means that' predicates, and hence neither under the disjoined predicate. Therefore the semantics is still incomplete. Obviously, the present argument can be repeated for any additional 'means that$_i$' predicate.

It is, however, possible to reach completeness by adding a countable infinity of truth predicates. That is, we have a infinite set of *means that* predicates

$$\{\text{`means that'}_0, \text{`means that'}_1, \text{`means that'}_2, \ldots\}$$

With this infinite set it holds that for any liar sentence s formed by a simple truth predicate, or by a finitely disjoined truth predicate, there is a simple truth predicate, formed by 'means that$_k$' for some k, that can be consistently applied to s and can represent within the language the intuitive truth of s. Then, if the language does not have infinite disjunctions, there is no highest truth predicate, and therefore no superliar sentence that cannot be handled in the language itself. The language is semantically complete in this sense, although there is no single predicate that internally expresses the meaningfulness property or the truth property of the language.[10]

However, leaving aside the lack of elegance, the proposal to have infinitely many *means that* predicates has two severe problems. First, since each of these predicates needs separate introduction, and a separately stated disquotation principle, it is in practice unreachable. Secondly, the semantic completeness of such a language is destroyed by the addition of new devices. For we can replace the infinite set by a single three-place predicate '... means that* ... at index ...', which takes the index as a third argument. For this we need names in the language of the indices, typically by means of numerals, i.e. expressions for the natural numbers.

In order to reduce also the number of disquotation principles to just one, we need something more. We need to be able to *quantify* into the index position. With this device we can state the general disquotation principle

(D*) $\forall p(\text{if } \exists n(\ulcorner q \urcorner \text{ means-that* } p \text{ at index } n),\ \text{then } p \text{ iff } q).$

However, with quantification into index position, we also have a highest, comprehensive truth predicate:

$$\exists p \exists n(\ldots \text{means-that* } p \text{ at index } n \;\&\; p)$$

which in turn allows us to form a superliar sentence

(54) $\neg\exists p \exists n((54) \text{ means that } p \text{ at index } n \;\&\; p)$

That this is a superliar is easily verified. For by familiar reasoning we can, for each index k, derive the conclusion

(55) $\neg\exists p((54) \text{ means that* } p \text{ at index } k)$

which then generalizes into

[10] Quine (1992, 89) suggests at one point having infinitely many truth predicates within the same language, as an alternative to Tarski's meta-language hierarchy. In Quine's proposal there is instead a hierarchy of truth predicates.

(56) $\neg\exists p \exists n((54)$ means that* p at index $n)$

from which we immediately infer

(57) $\neg\exists p \exists n((54)$ means that* p at index n & $p)$

i.e. (54). The semantics is again incomplete. Before turning to a discussion of the situation, I shall try to provide a model theory for a language L_1 with such a three-place *means that* predicate and quantification into index position.

5. Modeling extended internal semantics

To define the language L_1 we need to introduce (singular) index terms, both closed index terms and index variables.

Definition 5.1. The vocabulary of the language L_1 consists of

a. denumerably many sentence letters 's_1', 's_2', ...

b. denumerably many primitive singular sentence terms '(1)', '(2)', ...

c. denumerably many closed singular index terms 'i_1', 'i_2', ...

d. denumerably many propositional variables 'p_1', 'p_2', ...

e. denumerably many index variables 'x_1', 'x_2', ...

f. propositional logic constants '$\neg$' and '&'

g. quantifier symbol '$\exists$'

h. the three-place relational expression 'R'

i. the quote sign '$\ulcorner\urcorner$'

As before we need to define terms, formulas and sentences together.

Definition 5.2. The sets of well-formed sentence terms (wfsts) and well-formed formulas (wffs) of L_1 are simultaneously defined inductively as the smallest sets satisfying:

a. every sentence letter of L_1 is a wff

b. every propositional variable of L_1 is a wff

c. if ϕ and ψ are wffs, then $\phi \& \psi$ is a wff

d. if ϕ is a wff, then $\neg\phi$ is a wff

e. if ϕ is a wff and q a propositional variable, then $\exists q\phi$ is a wff

f. ϕ is a sentence iff ϕ is a wff without free variables

g. if ϕ is a sentence, then $\bar{\phi}$ is a wfst

h. if t is a wfst, ϕ a sentence, and i index term, then $R(t, \phi, i)$ is a wff

i. if ϕ is a wff and v an index variable, then $\exists v\phi$ is a wff

As before, we need restrictions against circularity in order to insure the consistency of the semantics. In the case of L_1, the restrictions need be more complex. To see this, consider the pair of sentences

(58) $\exists p(R((59), p, i_1) \;\&\; p)$

(59) $\neg\exists p(R((58), p, i_2) \;\&\; p)$

By the definition of self-dependence of section 3, both (58) and (59) are self-dependent. For (58) depends on the term '(59)' occurring in it, and this term depends on sentence (59) that it denotes, and (59) in turn depends on the term '(58)' occurring in it, and '(58)' depends on sentence (58) which it denotes. But inconsistency does not result from all internal meaningfulness assumptions. The assumptions

(60) a. $\exists p(R((59), p, i_2))$

b. $\neg\exists p(R((58), p, i_1))$

do jointly lead to contradiction, but the assumptions

(61) a. $\exists p(R((59), p, i_3))$

b. $\neg\exists p(R((58), p, i_4))$

do not. So to get the relevant dependence we have to relativize it to index, as follows:

Definition 5.3. Dependence:

a. A sentence s depends on a sentence s' with respect to index i if s contains a clause

$R(t, \ldots, t')$, where t denotes s and t' denotes i or t' is a bound variable.

b. If an expression s depends on a sentence s' with respect to some index and and s' depends on a sentence s'' with respect to index i, then s depends on s'' with respect to i.

c. An sentence s is *self-dependent* with respect to i iff s depends on itself with respect to i by clauses (a)–(b)

Definition 5.4. A model $\mathcal{M}$ for L_1 is a tuple $\langle I, O, V\rangle$, where I is an interpretation function denotations to simple terms and and sentences, O is a domain of indices, and V is a valuation function such that

a. Where t is simple sentence term [index term], $I(t)$ is a sentence of L_1 [$I(t) \in O$]

b. For a primitive term t, $V(t) = I(t)$

c. $V(\bar{s}) = s$

d. If s is a sentence letter, $\mathcal{M} \models s$ iff $I(s) = \mathrm{T}$

e. $\mathcal{M} \models s \& s'$ iff $\mathcal{M} \models s$ and $\mathcal{M} \models s'$

f. $\mathcal{M} \models \neg s$ iff not $\mathcal{M} \models s$

g. $\mathcal{M} \models R(t, s, i)$ iff $V(t) = s$ and s is not self-dependent with respect to i.

h. $\mathcal{M} \models \exists p \phi$ iff for some substitution instance $\phi(s/p)$, $\mathcal{M} \models \phi(s/p)$

i. $\mathcal{M} \models \exists x \phi$ iff for some substitution instance $\phi(i/x)$, $\mathcal{M}, g \models \phi(i/x)$

With this truth definition we can again verify that predicates like '$\exists p(R(t, \ldots, i)\ \&\ p)$' and '$\exists p \exists x(R(t, \ldots, x)\ \&\ p)$' behave as truth predicates. For if we have

(62) $\mathcal{M} \models \exists p \exists x(R(t, p, x)\ \&\ p)$

then by clauses h) and i) there is an instance $R(t, s, i)\ \&\ s$ such that

(63) $\mathcal{M} \models R(t, s, i)\ \&\ s$

which holds only if $V(t) = s$ and $\mathcal{M} \models s$. Conversely, if $V(t) = s$ and $\mathcal{M} \models s$ and s is not self-dependent with respect to i, then we have (63), and therefore by double generalization

(62).

The disquotation schema

(D!) $\mathcal{M} \models \neg\exists p \exists x(R(\bar{s}, p, x) \ \& \ \neg(\neg(p \& \neg s) \ \& \ \neg(s \& \neg p)))$

can similarly be seen to be valid. For a counterexample is a doubly existential sentence, and if a singular instance of such a sentence is true, then there is also a counterexample to the simpler disquotation schema (D+) of section 3, and we already saw that there isn't.

A superliar sentence like

(64) $\neg\exists p \exists x(R((64), p, x) \ \& \ p)$

gets evaluated as true, for (64) contains the clause $R((64), s, x)$, where 'x' is a bound variable, and '(64)' refers to (64) itself, and hence by definition 5.3, (64) is self-dependent with respect to any index. So the clause $R((64), p, x)$ will be evaluated as false for any value of 'p' and 'x'.

6. How to live with semantic incompleteness

We have seen that with the device of propositional quantification and unrestricted singular reference, the relational 'means that' expression generates liar sentences for a defined sentential truth predicate. We have also seen that with the restricted disquotation principle that belongs with this truth predicate, inconsistency can be blocked.

The price to be paid is a certain semantic, or expressive, incompleteness. A liar sentence, or a member of a liar chain of sentences, is intuitively true, since derivable by standard principles from a true sentence, is still counted as not expressing any proposition. This is not acceptable as a definitive situation.

It seems that we can avoid any particular incompleteness by way of the possibility of extending the internal semantic vocabulary with new *means that* predicates. We can do this by successively extending the domain of indices. But this leads to a dilemma, depending on whether we take the domain of indices to be extensible or not.

Suppose first that it is not. In that case there is a fixed total domain D of indices, whether they are ordinal numbers or something else. We can then introduce D as our quantification domain of indices. By means of quantifying over this domain we can produce the absolutely maximal truth predicate,

$$\neg\exists p \exists x \in D((\ldots \text{ means that } p, x) \;\&\; p)$$

(if we make the domain explicit) and by means of it we can produce the absolutely maximal liar

(65) $\neg\exists p(\exists x \in D)(((65) \text{ means that } p, x) \;\&\; p)$

As usual, we can show that (65) does not express any proposition, and also prove (65) itself. Since there is no way of extending the domain of indices, there is no way of capturing the intuitive judgment that (65) is meaningful after all. Instead, we have to stick with what we can express, and conclude precisely that (65) is not meaningful, or at least does not express a proposition. But if so, at least one of the inference rules by which (65) can be inferred from the acceptable

(66) $\neg\exists p(\exists x \in D)((65) \text{ means that } p, x)$

must be flawed, since leading from true premises to a meaningless conclusion. This might be conjunction elimination, existential elimination (with respect to propositional quantification), or negation introduction. However, that in itself just too implausible. This alternative does not seem acceptable.

The other horn of the dilemma is that if we take the domain of indices to be indefinitely extensible, we land in the problems of domain restrictions discussed in section 2.

It may seem that there is a way out of this dilemma, precisely by appeal to restriction of 'means that', as follows. We accept that there is a non-extensible total domain D of indices. What matters is what domain of indices the three-place 'means that' is *defined* for. We can *choose* to extend its definition to cover a new, more extensive sub-domain of indices, i.e. a new, more extensive sub-domain of D. Deciding to make it defined for D itself is incoherent, as the dilemma above shows, and if 'means that' *isn't* defined for the total domain D, (66) is not itself a meaningful sentence. As long as we only define it for sub-domains (built up from below), we can at any one point extend the index domain for which it is defined at that point.

Then we have to say about the liar sentence

(7) $\neg\exists p((7) \text{ means that } p \ \& \ p)$

that it there is no proposition such that (7) expresses that proposition. I *can* say that there is an admissible extension i of 'means that' such that the sentence

(67) $\exists p((7) \text{ means that}_i \ p \ \& \ p)$

can be made assertible. However, I cannot say *what will be asserted* by means of (67) until the extension is actually made. And thirdly, we still have to say, strictly speaking, that apparently acceptable rules of inference in some special cases lead from acceptable premises to meaningless conclusions.

On the other hand, the following would hold, if the solution were acceptable: "as long as we don't decide make our 'means that' defined for all indices, there is no *definitive* limit to our ability to extend our internal semantics. This means, among other things, that the rules of inference in question can be seen as valid in a somewhat extended sense: from acceptable premises they lead to conclusions that are either actually or *potentially* acceptable, never to a conclusion that is definitively flawed. Moreover, on this approach, we can actually state that much --- as I am doing now --- without transcending the expressive restrictions that we now have. That is, I can say that, without transcending the expressive limitations of my present language, that the interpretation of 'means that' is indefinitely acceptably extensible."[11]

[11]It may be objected that because of compositionality, '$\xi_1 \text{means that*} \ \xi_2 \text{ at stage } \xi_3$' is defined for all argument expressions anyway, including all index terms that we may have, so that it is simply a violation of compositionality to claim that this predicate isn't well-defined for some fillers of the third argument place. But that is a misconception. If the *concept* expressed (at a certain stage) by this three-place predicate isn't applicable to triples with a certain index as the third element, then the function mapping meanings of parts on meanings of wholes is not defined for a combination of arguments that includes this concept and that index either. And then the sentence you get by filling the third argument place by term referring to the index is not meaningful.

When we extend the interpretation of 'means that', we don't extend the concept it expresses (I take concepts to be abstract, immutable entities) but start associating the expression with a different and somewhat wider concept. We can indeed *predict* what concept our 'means that' will be associated with after any particular extension, but this is not enough for generating inconsistency. For generating inconsistency we need an

If this apparent way out were good, it would be possible to quantify over admissible extensions of the semantics of the 'means that' predicate without introducing new liar sentences. The idea would be that as long as an admissible extension is not in fact *made*, our use of 'means that' would not have a meaning that includes the extension, and hence could not be exploited for constructing a new liar.

The problem with the suggestion is that precisely in giving a general description of the procedure of revising and extending the 'means that' predicate, we make use of expressive resources that transcend the expressive limitations we wanted to impose. For the talk of extending the interpretation of 'means that' at a time t involves talk of what 'means that' *will mean* at t. We have then tacitly accepted to use 'means that', or 'means', with an extra time variable ranging over all times. That is, we allow ourselves to talk about what meaning something will have at some time without being able to express that meaning. That allows us to form a new three-place predicate R' holding between a sentence s, a proposition p and an index i just if at some time or other, 'means that' will express a relation that holds between s and p:

$$R'(s, p, i) \leftrightarrow \exists t(\text{'means that' means } F \text{ at } t \ \& \ F(s, p, i))$$

And this, in turn, gives us the possibility of constructing a new liar sentence:

(68) $\neg\exists p \exists n(R'((68), p, n) \ \& \ p))$

(68) says that there will never be a time t, a proposition p and an index i such that at t 'means that' expresses a relation holding between (68), p and i, and such that p. Together with a corresponding disquotation principle for 'R'', this is a liar sentence that cannot be resolved by extending 'means that', for it already quantifies over such extensions. Nor can we reject the disquotation principle

(DR) $\forall p \forall n(\text{if } R'(\ulcorner q \urcorner p, n), \text{ then } p \text{ iff } q)$

for the negation of any meaningful instance of (DR) is inconsistent.

expression in the language that already *has* that meaning. This expression is needed for stating the disquotation principle that is required for deriving the contradiction.

The conclusion, then, is pretty dark. We cannot even describe how we can successively cope with liar sentences while being able to express our intuitive judgments about meaningfulness. In a sense we *can* successively cope with them, but as soon as we allow ourselves the resources to say that much, we lose the ability. This is, in a way, old news from Tarski, just spelled out somewhat differently.

Peter Pagin
Department of Philosophy
Stockholm University
106 91 Stockholm
Sweden
email: peter.pagin@philosophy.su.se

References

Burge, T., (1979), `Semantical paradox', *Journal of Philosophy* 76:179--98.

Glanzberg, M., (2004), `A contextaul-hierarchical approach to truth and the liar paradox', *Journal of Philosophical Logic* 33:27--88.

Hill, C. S., (2002), *Truth and World. An Austere Portrayal of Truth, Reference and Semantic Correspondence*, Cambridge University Press, Cambridge.

Kaplan, D., (1994), `A problem in possible world semantics', in W. Sinnott-Armstrong and alt. (eds.), *Modality, Morality and Belief. Essays in Honor of Ruth Barcan Marcus*, Cambridge University Press, Cambridge.

Lindstrom, S., (2003a), `Possible worlds semantics and the liar', in A. Rojszczak, J. Cachro and K. Kurszewski (eds.), *Philosophical Dimenson of Logic and Science*, volume 320 of *Synthese Library*, Kluwer Academic Publishers, Dordrecht.

Lindstrom, S., (2003b), `Frege's paradise and the paradoxes', in K. Segerberg and R. Sliwinski (eds.), *A Philosophical Smorgasbord: Essays on Action, Truth, and Other Things in Honour of Fredrick Stoutland*, volume 52 of *Uppsala Philosophical Studies*, Uppsala University, Uppsala.

Parsons, C., (1974), `The liar paradox', *Journal of Philosophical Logic* 2:381--412.

Quine, W. V. O., (1992), *Pursuit of Truth*, Harvard University Press, Cambridge, Mass., 2nd edn.

Russell, B., (1908), `Mathematical logic as based on the theory of types', *American Journal of Mathematics* 30:22--262.

Simmons, K., (1993), *Universality and the Liar. An Essay on Truth and the Diagonal Argument*, Cambridge University Press, Cambridge.

Williamson, T., (2003), `Everything', *Philosophical Perspectives* 17:415--65.