Notre Dame Journal of Formal Logic
Volume 27, Number 1, January 1986

Analyticity and Analytical Truth

ERMANNO BENCIVENGA

In the literature, there are *two* distinct characterizations of analyticity. One is purely syntactical: it refers to the structure of a sentence only. Limiting our attention for the moment to subject-predicate sentences, a first approximation to such a characterization can be given by the following variant of the classical Kantian definition:

(1) A sentence of the form 'the P is (a) Q' is analytic if P is a predicate that (in some sense to be specified) *includes* Q.

The second characterization brings semantics, and truth in particular, into the picture:

(2) A sentence is analytic if it is *true* by virtue of its form alone.

There are reasons to think that the two characterizations are in conflict. For consider sentences like

(3) The square that is not a square is a square that is not a square.
(4) The winged horse that exists is a winged horse that exists.

Whatever construal you give of 'includes' in (1), both (3) and (4) are analytic in the sense of (1). Then, since trivially (3) and (4) have the form they have, according to (2) they must be true. But for most people, (3) and (4) are *not* true. So if we assume that (1) and (2) are both characterizations of analyticity, we (apparently) reach an absurd conclusion.

People have had two basic reactions to this problem. Some have insisted that (3) and (4) *are* true, that is, that the above conclusion is *not* absurd. The results of this attitude have been various kinds of Meinongian or dialectical logics, committed to either the claim that there are nonexistent objects or the claim that reality is contradictory (or both).[1] The difficulty with such results is that they saddle logic with debatable (though in themselves perfectly respectable) metaphysical theses: if these commitments are accepted, it would seem to follow that one cannot always argue rationally with people who do not accept such theses (in fact, that one cannot always understand what they are saying).

Received June 28, 1984; revised December 3, 1984

On the other hand, some people have simply given up on the analyticity of (3) and (4), and tried to weaken (1) in some such form as the following:

(5) 'If the P exists, then the P is Q' is analytic if P includes Q.[2]

Note however that adopting this attitude would prevent us from attributing analyticity to sentences like

(6) The President of the United States is a President of the United States,

too, and hence from making what looks like a perfectly reasonable distinction between (the structure of) (6) and

(7) The President of the United States is a tall man.

Furthermore, the attitude in question would also make us unable to single out the structural similarity of (6) *and (3)*, which evidently subsists in spite of (3)'s contradictoriness.

So to put it simply, one choice extends logic too much, and one restricts it too much. One ends up having logic cover a lot of metaphysics as well, and one prevents logic from studying interesting structural properties of sentences. Here I would like to propose a third alternative, which has neither of the above negative consequences. The first basic step of this alternative consists in reading (1) and (2) as characterizations of *distinct* properties, that I will call *analyticity* and *analytical truth*, respectively.[3]

Of course, introducing new terminology is not enough to address a philosophical problem. Here however I am in a position to do more than that: I can present a formal semantics in which the two (informal) characterizations in (1) and (2) are given formal counterparts, that is, turn out to correspond to two *formally identifiable* (and distinct) classes of sentences.[4] The semantics is one that I developed in a different context, and for different reasons (see [1]). The fact that it has this additional application is evidence that it captures some basic logical intuitions.

A brief sketch of the semantics is as follows.[5] The basic unit is a *free model* (for a language L) that is, a pair $\langle D, f \rangle$, where D (the domain) is a set (possibly empty) and f (the function of interpretation) is total on the predicates but partial on the constants of L. A valuation V_M of the terms and sentences of L is defined for each model M, in such a way that

(a) when all the terms occurring in an atomic sentence A are denoting in M, $V_M(A)$ is computed as usual
(b) with two exceptions, when some of the terms occurring in (an atomic) A are nondenoting in M, $V_M(A)$ is not defined (the two exceptions are sentences of the form $\mathrm{E}!a$ and $a = b$, which are false when b is denoting and a is not)
(c) when there is a unique entity in M satisfying the formula B, $V_M(\iota xB)$ is that entity, and when there is no such unique entity, $V_M(\iota xB)$ is not defined
(d) complex sentences are evaluated as usual if none of their components are truth-valueless, and left truth-valueless otherwise.

Now suppose that a sentence A has no truth-value in V_M, and consider another free model M' that is a restriction of an extension of M. For each such M' we define a valuation $V_{M'(M)}$ for *M' from the point of view of M*, which assigns to each well-formed part B of A the truth-value B has in V_M if B has a truth-value there, and otherwise assigns to B the truth-value it has in $V_{M'}$ (if any).[6] It is possible that, for some such M', $V_{M'(M)}(A)$ is defined (that is, A has a truth-value in M' from the point of view of M); we call each M' for which this is the case an *A-world for M*. If there is no A-world for M, A remains truth-valueless in M. If there are A-worlds for M, we consider the supervaluation S_M on all the $V_{M'(M)}$ (where M' is an A-world for M), and we let these supervaluations be the admissible valuations of our semantics, so that A is true (false) in a free model M if and only if, *when looked at from the point of view of M*, A is true (false) in all A-worlds for M, and has no truth-value in M otherwise.

The following formulas are logically true in this semantics:

(8) $P(\imath x Px)$
(9) $P(\imath x(Px \,\&\, Qx))$

(where P is a primitive predicate; see below). On the other hand, the following formulas are *not* logically true in this semantics:

(10) $P(\imath x(Px \,\&\, {\sim}Px))$
(11) $P(\imath x(Px \,\&\, {\sim}Px)) \,\&\, {\sim}P(\imath x(Px \,\&\, {\sim}Px))$
(12) $\mathrm{E!}(\imath x \mathrm{E!}x)$
(13) $\mathrm{E!}(\imath x(\mathrm{E!}x \,\&\, Px))$.

More precisely, (12) is falsified by every free model not containing exactly one entity, (13) is falsified by every free model not containing exactly one (thing that is) P, and (10)–(11) are *essentially truth-valueless*, that is, have no truth-value in any free model.

Consideration of (10)–(13) above indicates that the most natural paraphrases of sentences like (3) or (4) are not logically true. And this is promising, since it is also natural to regard logical truth (that is, truth in all possible worlds) as a formal analogue of *analytical truth*. On the other hand, consideration of the logical truth of (8) and (9) suggests that we can probably use this semantics to cash out in formal terms the similarity between (3)–(4) and (6). For (3) and (4) are, just like (6) (but in contrast, for example, with (7)) *instances of logically true formulas*. To begin to make all of this precise, consider the following two definitions:

(14) Let A be a formula of L. Let $P_1, \ldots, P_n$ be predicates occurring in A, let $a_1, \ldots, a_m$ be constants occurring in A, and let $x_1, \ldots, x_p$ be variables occurring in A. Let each P_i be of degree d_i. Let $B_1, \ldots, B_n$ be formulas of L, and let each B_i contain exactly d_i free variables. Let $t_1, \ldots, t_m$ be closed terms of L. Let $s_1, \ldots, s_p$ be terms of L, where each s_i contains free exactly the variable x_i. Let the formula C result from A by substituting each B_i for the corresponding P_i, each t_i for the corresponding a_i, and each s_i for the corresponding x_i. Then C is a *substitution instance of A*.[7]

(15) A formula of L is *pseudologically true* if it is a substitution instance of a logically true formula.

Pseudological truth is going to be our formal analogue of analyticity, just as logical truth is of analytical truth. To apply this formal analogue to English sentences, we need to establish a correspondence between them and formulas of L. Simple paraphrase will not be enough, since both (8) and (11) are paraphrases of, for example, (3). What we want is a paraphrase that captures as much of the logical structure of an English sentence as is possible given the expressive means available in L. Together, the following two definitions achieve that.

(16) If A is a substitution instance of B, and A is longer than B, then A is a *proper substitution instance* of B.

(17) A formula A (of L) is a *maximal paraphrase* of an English sentence B if (a) A is a paraphrase of B, and (b) no proper substitution instance of A is a paraphrase of B.[8]

So finally we can agree that

(18) An English sentence is *analytic* if a maximal paraphrase of it (in L) is pseudologically true (or, alternatively, if any paraphrase of it is logically true).

(19) An English sentence is *analytically true* if a maximal paraphrase of it (in L) is logically true.

The notions defined in (18)–(19) do full justice to the parallels and distinctions we want to draw. For, though they make neither (3) nor (4) analytically true, they make them both analytic; (7), on the other hand, is arguably not analytic.[9] And in general, though the relation between (18) and (1) is not immediately evident—that is, though (18) does not result from a direct attempt at making it precise what 'includes' means—all the sentences that people (for example Kant) usually regarded as analytic on the basis of (1) turn out (in view of the logical truth of (8) and (9)) to be analytic in the sense of (18).

It is worth mentioning in closing that this result is made possible by a feature of the semantics in question that some may find objectionable, that is, the fact that the semantics does not admit the *rule of substitution*.[10] If this rule were admissible, then the logical truth of any formula A would entail the logical truth of all substitution instances of A, and pseudological truth would collapse into logical truth. Of course, the admissibility of the rule of substitution in a semantics makes it much easier to operate with that semantics, but in many cases (and the present one is among them) technical simplicity is obtained at the expense of expressive power.[11]

NOTES

1. See for example Routley's [7]. Parsons, on the other hand, is aware that his Meinongian theory (as presented in [6]) is a metaphysical and not a logical one.

2. Examples can be found (implicit) in Lambert's [5] and in [9] by van Fraassen and Lambert. Russell's position, too, falls in this general category, but it also contains the additional assumption that the surface structure of (3)–(4) (or, for that matter,

of (6)–(7)) is delusive (and faces the additional philosophical problem of justifying this assumption).

3. Another author who makes this distinction, and who also identifies several senses of analyticity, is Hintikka (see his [4], especially p. 124). But the intuitions behind Hintikka's treatment, and the purposes of that treatment, are quite different from mine here.

4. Even though the motivations for making the distinction above arise primarily from a consideration of descriptions, the formal semantics will allow for completely *general* characterizations. In particular, definition (15) below is a (formal) generalization of (1).

5. In this sketch, I have modified some of the terminology of the original paper. But all differences are self-explanatory.

6. A minor technical detail, which is significant in the original context but not here: in evaluating binary connectives, $V_{M'(M)}$ follows the strong Kleene truth-tables whereas V_M follows the weak ones.

7. This definition is not completely explicit, since it does not say what it is to substitute a formula for a predicate. If the reader is interested in a more technical formulation, he can find it for example in Church's [3] especially pp. 192–193.

8. Note that I give no formal definition of the notion of a paraphrase. But this was to be expected, since paraphrasing establishes a connection between a formal language and a natural, *in*formal one. Therefore, the notion of a paraphrase can only be understood at an intuitive, informal level (the level at which we conduct most introductory logic courses, before people begin to speak in symbolese).

9. To argue (informally, see note 8) for this, it is essential to note that the set of possible paraphrases of (7) into L can be exhausted after a finite search.

10. The semantics discussed here is not the only one that has this feature. Primary modal semantics (that is, the semantics that interprets modal logic on a single structure, containing *all* possible worlds) is another example. I discuss the failure of the rule of substitution in primary modal semantics in [2].

11. To mention only one other case in which this is true, think of the Strawson–Lambert–van Fraassen treatment of existential presuppositions (as presented for example in van Fraassen's [8]). This treatment is only possible because the semantics in question is not bivalent, but of course bivalence allows important simplifications at both the theoretical and the metatheoretical level.

REFERENCES

[1] Bencivenga, Ermanno, "Free semantics for definite descriptions," *Logique et Analyse*, vol. 92 (1980), pp. 393–405.

[2] Bencivenga, Ermanno, "On secondary semantics for logical modalities," *Pacific Philosophical Quarterly*, vol. 62 (1981), pp. 88–94.

[3] Church, Alonzo, *Introduction to Mathematical Logic*, Princeton University Press, Princeton, New Jersey, 1956.

[4] Hintikka, Jaakko, *Logic, Language-Games, and Information*, Clarendon Press, Oxford, 1973.

[5] Lambert, Karel, "Notes on E!III: A theory of descriptions," *Philosophical Studies*, vol. 13 (1962), pp. 51–59.

[6] Parsons, Terence, *Nonexistent Objects*, Yale University Press, New Haven, Connecticut, 1980.

[7] Routley, Richard, *Exploring Meinong's Jungle and Beyond*, Australian National University, Canberra, 1980.

[8] van Fraassen, Bas C., "Presupposition, implication, and self-reference," *Journal of Philosophy*, vol. 69 (1968), pp. 136–152.

[9] van Fraassen, Bas C. and Karel Lambert, "Free description theory," *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, vol. 13 (1967), pp. 225–240.

Department of Philosophy
University of California, Irvine
Irvine, California 92717