Chapter XVII

WHAT IS RUSSELL'S THEORY OF DESCRIPTIONS?[1]

David Kaplan
University of California, Los Angeles

Russell expounded his theory of descriptions in a number of places, but perhaps the best known source is his 1905 article, "On Denoting" [1]. I think it may still be fruitful to discuss the doctrine of that article since some readers may disagree as to its main point.

Theories of descriptions concern the analysis of sentences containing definite descriptions. For example, "The present queen of England is shapely" or "The least prime number is even." Let us refer to the description itself as proper or improper according as there is or is not a unique object described. Thus the descriptions "The present queen of England" and "The least prime number" are both proper since they uniquely describe Elizabeth Windsor and two, respectively. But the descriptions "The present king of France" and "The author of *Principia Mathematica*" are improper since the first describes nothing and the second describes Russell and Whitehead equally well. The difficulties involved in the analysis of sentences containing descriptions (I will say "description" as short for "definite description") are most apparent in connection with improper descriptions. This should not be surprising, since improper descriptions are rarely used knowingly, and thus usage does not provide a clear guide. On even the most elementary question of analysis, the truth value of such sentences, we find disagreements. Consider the sentence: "The present king of France is bald." According to Russell's theory, it is false. According to the "chosen-object" theory of Frege[2], elaborated by Carnap[3], in which all improper descriptions are treated as if they uniquely described some previously chosen

[1]This paper has benefited from a number of sources, including a seminar at Princeton, a colloquium at Cornell, discussions with Montgomery Furth, and National Science Foundation Grant GP-4594.

W. Yourgrau et al. (eds.), *Physics, Logic, and History*
© Plenum Press, New York 1970

object, the sentence is true (taking Yul Brynner as chosen object). According to the "truth-value-gap" theory of Frege,[2] elaborated by Strawson[4], in which improper descriptions are treated as having meaning but describing nothing, and sentences containing such descriptions are treated as themselves meaningful but having no truth value, the sentence is neither true nor false. I am aware of no theory according to which the sentence is both true and false, though no doubt such a theory has been or will be proposed.

This much is well known about Russell's theory: He takes the propriety of the description to be a part of the *content* of certain sentences containing descriptions and thus counts them false if they contain *improper* descriptions. More specifically, he claims that a paradigm sentence of the form "The such-and-such is so-and-so" is equivalent to "One and only one thing is a such-and-such, and that one is so-and-so." In symbols:

$$\text{“}F\imath xGx\text{” is equivalent to “}(\exists x)\,((y)\,(Gy \equiv y = x)\ \&\ Fx)\text{.”} \qquad (1)$$

This analysis does not provide a unique understanding of sentences of the form "The such-and-such is not this-and-that," which may be treated as equivalent to "One and only one thing is a such-and-such and that one is not this-and-that," thus assimilating "not this-and-that" to the "so-and-so" of the paradigm; but the given form may also be treated as equivalent to "It is not the case that one and only one thing is such-and-such and that one is this-and-that," thus applying the paradigm to "the such-and-such is this-and-that" and understanding the given form as its negation. This ambiguity is regarded by Russell as a simple scope problem on a par with the party invitation which reads, "Bring your wife or come stag and have a good time." And he introduces the terminology of *primary* and *secondary* (and by natural extension *tertiary*, *quaternary*, and so forth) occurrences to indicate the intended scope of the description in sentences of the given form. (Later, in *Principia Mathematica* [5], he introduces the more technically satisfactory device of scope indicators.)

These two features then, (i) that the sentence "The present king of France is bald" is taken to be equivalent to "One and only one thing is a present king of France and that one is bald," and (ii) that this leads to scope problems in the case of "The present king of France is not bald" which comes out true if the description is given secondary occurrence but false if the description is given primary occurrence, I take

[2] "Über Sinn und Bedeutung" [2]. See especially the discussion of sentences containing "Odysseus."

to be well known. And, possibly because they provide a convenient means of comparison with other theories of descriptions, they have sometimes been thought to constitute the distinctive feature in Russell's doctrine. But this, I believe, is not so. The peculiar and interesting feature is his position on denoting. For he claims that, in contrast to such proper names as "Elizabeth Windsor" which denote a certain person, a description like "The present queen of England" doesn't *really* denote anything. If the description is proper you may want to speak *as if* it denoted the unique thing described, and that won't get you in any difficulties (if you avoid oblique contexts), but it would be misleading. For it would lead you to believe that the two sentences:

"Elizabeth Windsor is shapely," (2)

"The present queen of England is shapely" (3)

have the same logical form, namely subject-predicate form, with "Elizabeth Windsor" in one case and "the present queen of England" in the other case as subjects, both denoting the same individual, and "is shapely" as predicate, denoting, say, some class of individuals. Thus in both cases the sentence is true if, and only if, the given individual is a member of the given class. But nothing could be further from the Russellian truth. Sentences (2) and (3) do not have the same logical form at all, for if (3) were of subject-predicate form, so would

"The present king of France is shapely" (4)

be, by parity of form (as Russell would say). But according to Russell the truth conditions for (4) have nothing to do with any given individual being shapely.

And now, I believe, we are at the heart of the matter. Russell's article "On Denoting" is not about a theory of descriptions comparable to Frege-Carnap or Frege-Strawson. Russell's article is about logical form, and is in the tradition of those philosophers who have warned us of the dangers of confusing the grammatical form of a sentence in ordinary language with its logical form. Such philosophers have often sought to construct a *logically perfect language* in which grammatical and logical form would always coincide.

I am dissatisfied with Russell's theory about the logical form of sentences containing descriptions, and I will try to indicate the features I find unsatisfactory by comparison with an alogous theory which I call "Russell's theory of indefinite descriptions."[3] This theory is con-

[3]Russell's theory of indefinite descriptions is most clearly adumbrated in [15].

cerned with the analysis of sentences containing *in*definite descriptions. For example:

"A senator from New York is supporting Rockefeller." (5)

Now (5) certainly has subject-predicate grammatical form in English, but if you feel that its logical form is the same as

"Jacob Javits is supporting Rockefeller," (6)

you can quickly disabuse yourself by comparing:

"A senator from New York is supporting Rockefeller, and a senator from New York is not supporting Rockefeller," (7)

with

"Jacob Javits is supporting Rockefeller, and Jacob Javits is not supporting Rockefeller." (8)

Sentence (8) is a contradiction, but (7) is true. In fact, isn't it obvious that indefinite descriptions do not even purport to denote a unique object as names do? Accordingly, Russell's theory of indefinite descriptions asserts that the logical form of a paradigm sentence like "A such-and-such is so-and-so" is represented by the equivalent sentence "Something is both a such-and-such and so-and-so." In symbols:

"$F\alpha xGx$" is equivalent to "$(\exists x)\,(Gx\ \&\ Fx)$". (9)

Note that this analysis does not provide a unique understanding of certain compound sentences containing indefinite descriptions, for example:

"A girl danced with every boy." (10)
In symbols: $(y)\,(By \supset D(\alpha xGx,y))$.

Depending on whether the indefinite description is taken as having primary or secondary occurrence, the sentence will be equivalent to either

"Some girl is such that she danced with all the boys," (11)
in symbols: $(\exists x)\,(Gx\&\ (y)\,(y)\,(By \supset D(x,y)))$

or

"Each boy is such that some girl or other danced with him", (12)
in symbols: $(y)\,(By \supset (\exists x)\,(Gx\ \&\ D(x,y)))$

There are further parallels between Russell's theory of definite descriptions and Russell's theory of indefinite descriptions, but I think the point is made. The point is that Russell regarded definite descriptions exactly as we would regard indefinite descriptions. Grammatically, at least from the point of view of what is now called surface grammar, indefinite descriptions are terms and they function like proper names; but sentences which contain indefinite descriptions and appear to have subject-predicate form should be treated as idioms and expanded as in the paradigm.[4]

Russell's theory of indefinite descriptions seems to me both correct and natural. In fact, the analysis fits indefinite descriptions so perfectly that any disanalogies between definite and indefinite descriptions throw suspicion on Russell's theory of definite descriptions. Further, I believe that the leading ideas of Russell's theory of definite descriptions are more clearly seen in connection with Russell's theory of indefinite descriptions, and thus questions about the former might be more easily answered in connection with the latter.

With respect to disanalogies between definite and indefinite descriptions, I will just mention two ways in which Russell (and everyone else, I suppose) provides differential treatment. First, Russell invents a symbolic notation for definite descriptions and introduces them into the language of *Principia Mathematica*. So far as I know, neither Russell nor anyone else has ever given serious consideration to the introduction of indefinite descriptions into any formalism. Not that it could not be done. The foregoing brief remarks on Russell's theory of indefinite descriptions make it clear exactly how to do it. It is even clear how scope indicators could be introduced. It is just that nobody would think it worth doing. Why? Because such a notation, rather than providing a useful and succinct means of expression for investigating logical relations, would tend to obscure the logical form of the sentence and obfuscate the issues in question. This, of course, is exactly what definite descriptions of English are said (by Russell), to do but still he introduces them into *Principia Mathematica*.

The second respect in which Russell treats definite and indefinite descriptions differently is in connection with his offhand remark that one *might* speak of proper definite descriptions *as if* they denoted the unique individual having the property in question, although strictly speaking this would be incorrect. We certainly do want to treat proper descriptions in this way, and every other theory of definite descriptions with which I am familiar does so. Are there any cases (let alone the central cases) in which

[4]Compare "It is snowing," which also appears to have subject-predicate form.

it is natural to treat an indefinite description like a proper name, that is, as denoting an individual? Russell never suggests so.[5]

I wish now to raise the question of how to regard the fundamental equivalences (1) and (9) of the two theories. Russell called them *contextual definitions*. But what is a contextual definition? Ordinarily we think of definitions as being either stipulative or explicative (in an older terminology, nominal or real). That is, either a new expression is introduced and assigned the meaning of a phrase whose meaning is antecedently known, or else an old expression is given a more precise, or in some other way slightly adjusted, meaning in terms of some antecedently understood phrase. In the case in question, the new expression to be introduced is the definite or indefinite description operator. But what meaning is given to it, or even to the full description? None! For the central thesis of Russell's theory is that this phrase *has* no meaning in isolation.

Another notion closely related to that of definition which is, I believe, somewhat more appropriate to Russell's contextual definitions, is that of *abbreviation*. In an abbreviation a new expression is introduced to *stand for* an old phrase. But the phrase so abbreviated is not required to be meaningful; it may be any combination of signs. Abbreviation is purely a matter of syntax. Whereas only well-formed expressions can serve as definiens, any expression can be abbreviated. To understand the meaning of an expression in terms of the meanings of its components, we must first expand it into unabbreviated form (except, of course, insofar as the abbreviations are also definitions).

I prefer still another understanding of (1) and (9). We may treat them as rules for translating ordinary, logically imperfect language into a logically perfect symbolism. According to this conception the symbolic descriptions "αxFx" and "ιxFx" would be understood as not occurring in the perfect language at all, not even as abbreviations. They would appear only at a transitional stage in the translation process as an aid in representing the surface grammar of ordinary language. This seems to me a natural understanding of (9), since, as remarked above, no one would seriously consider introducing indefinite descriptions into a logically perfect language.[6] If Russell really took his doctrine seriously and were willing to completely abjure the 'misleading' surface grammar of definite descrip-

[5]At the beginning of "On Denoting," Russell suggests as an initial understanding of indefinite descriptions that " 'a man' denotes. . .an ambiguous man." But he quickly rejects this idea.

[6]Hilbert's ϵ-operator does not produce an indefinite description in the sense herein discussed.

tions, he should be willing to accept the present understanding of (1). His use of definite descriptions in *Principia Mathematica* indicates to me a lingering ambivalence.

Although I will not now attempt to make the notion of a logically perfect language absolutely precise, I would like to clarify it somewhat. The intuitive idea is that the logical form of a sentence should mirror its grammatical form. The grammar of a language is assumed to be given in terms of certain grammatical categories such as term, formula, two-place predicate, etc. Each atomic expression is assigned to some such category, and *formation rules* are given which tell us how we can form compounds of a given grammatical category from components of certain grammatical categories. The grammatical form of an expression is then determined by the formation rules. An expression is grammatically correct if it can be 'constructed' from grammatically simple components in accordance with the formation rules. Such a construction assigns a grammatical structure, or form, to the expression. To parse a sentence is to exhibit its grammatical form. Just as grammatical properties and relations, such as being a noun clause or being the subject of a given sentence, depend on the grammatical form of the expression in question, so logical properties and relations, such as being valid or being a logical consequence of a given sentence, depend on the logical form of the expressions in question. Logical form is determined by the *evaluation rules* of the language. These rules tell us how to 'construct' the semantical value of an expression in terms of the values of its logically simple components. (We here take the semantical value to be what Carnap calls "the extension," that is: a truth value for sentences, an individual for names, a class of individuals for one-place predicates, and so on.) Such a construction of the truth value of a sentence exhibits the logical structure, or form, of the sentence in a way analogous to that in which parsing a sentence exhibits its grammatical form. As shown by Tarski[6], the notions of validity and logical consequence can be given in terms of such constructions.[7] In ordinary language, replacements which do not change the apparent grammatical form of sentences, for example, replacing a proper name with "someone," may well introduce or obliterate relations of logical consequence between the affected sentences, thus indicating a change in logical form. This point, that sentences with the same apparent grammatical form can have different logical forms, is illustrated by (5)-(8). In a logically perfect language the logical form of an expression must always mirror the grammatical form. Therefore, for logical perfection we

[7]What I call evaluation rules are clauses in Tarski's definition of "satisfaction."

require that the logically simple expressions coincide with the grammatically simple (but well-formed) expressions, and that to every formation rule there corresponds a unique evaluation rule, such that any compound formed by applying the formation rule to given components is evaluated by applying the corresponding evaluation rule to the values of the components. This has the desired result that the semantical evaluation of an expression exactly recapitulates its grammatical construction.[8]

Given the grammar of a language, one semantical treatment may make it logically perfect and another not. Take for example a sentential language which contains: (1) the atomic expressions "P_1", "P_2", "P_3", etc., all belonging to the grammatical category *sentence*, and (2) three formation rules which allow us to form the compound sentences $\ulcorner(\Phi \supset \Psi)\urcorner$, $\ulcorner\sim \Phi\urcorner$ $\ulcorner(\Phi \equiv \Psi)\urcorner$ from any component sentences Φ and Ψ. Now consider two different methods for assigning values to the sentences. Method I consists of (1) assigning a truth value to each atomic sentence in accordance with some given interpretation of the atomic sentences, and (2) for each formation rule using a corresponding truth function to evaluate the truth value of the compound in terms of the truth values of its immediate components. Method I is the standard semantical analysis of such a language. Method II agrees with Method I for the atomic sentences, but provides two stages for the analysis of compounds, (a) 'eliminate' all biconditionals from compound sentences by replacing sub-sentences $\ulcorner(\Phi \equiv \Psi)\urcorner$ by $\ulcorner\sim((\Phi \supset \Psi) \supset \sim(\Psi \supset \Phi))\urcorner$, (b) evaluate the result as in Method I.[9] I call the language which incorporates Method I logically perfect, but the detour from grammatical form involved in Method II renders the language incorporating that method logically imperfect. This obtains in spite of the fact that the two methods assign the same values to all sentences.

Let us assume that the grammar of Russell's language distinguishes *term* and *formula* and contains among its formation rules: All variables are terms; all individual constants are terms; if Φ is a formula and v is a variable, $\imath v\Phi$ is a term; if τ is any term (a variable, individual constant, or definite description), $\ulcorner\tau$ is bald$\urcorner$ is a formula; plus the usual formation rules for identity, quantifiers, sentential connectives, etc.[10] Following Russell's informal remarks, we will understand his semantics as involving first the elimination of all descriptions by means of (1). In accordance with

[8]This relationship between formation rules and evaluation rules is developed in somewhat greater detail in Chapter 1 of my dissertation [7].

[9]It is important for the point of the example that all three formation rules are understood as primitive, and thus that biconditionals are *not* thought of as defined expressions.

[10]At this point we revert to the theory of definite descriptions and leave indefinite descriptions aside as an instructive amusement.

the conventions for dropping scope indicators in *Principia Mathematica*, we take the scope always to be the smallest possible. Such a semantical analysis will of course make (1) true, and in a trivial way. But it makes the language logically imperfect.[11]

This brings us to another question. Does acceptance of the equivalence (1) commit us to Russell's analysis of the logical form of sentences containing definite descriptions? The answer is "No." At least two different (but closely related) analyses can make the language logically perfect. We can follow the method of Frege-Strawson in claiming that improper descriptions simply don't denote, but in the Frege-Strawson evaluation rules for atomic formulas:

(i) $\ulcorner\tau$ is bald$\urcorner$ is true if and only if τ denotes something which is bald;
(ii) $\ulcorner\tau$ is bald$\urcorner$ is false if and only if τ denotes something which is not bald

(where τ may be any term: a variable, individual constant, or definite description); retain (i) and replace (ii) with:

(ii′) $\ulcorner\tau$ is bald$\urcorner$ is false if and only if it is not true.

Alternatively, we can follow the method of Frege-Carnap in stipulating that some previously chosen entity will be taken as the common denotatum of all improper descriptions, but put this entity *outside* the domain of discourse (possibly by just letting the chosen entity be the domain itself). According to this method we would retain the evaluation rule (ii) and replace (i) with:

(i′) $\ulcorner\tau$ is bald$\urcorner$ is true if and only if τ denotes something which is bald *and is within the domain of discourse*.

Note that in the unmodified Frege-Carnap theory the italicized phrase is otiose.

[11] It should be noted here that I assume the grammar to be given in the way described above, essentially what is now called an *immediate constituent phrase structure grammar*. It is not implausible to regard Russell's contextual definitions as providing his language with a *transformational phrase structure grammar*. Since Russell's implied evaluation rule for sentences containing definite descriptions, like that of our Method II, might plausibly be called a *transformational evaluation rule*, we might consider this mirroring of syntactical structure a kind of perfection. But in-so-far as there is a natural tendency to expect a language to exhibit the simpler immediate constituent form, a language whose grammar is described by transformational means might, on that account, be considered imperfect.

On both the modified Frege-Strawson analysis and the modified Frege-Carnap analysis the equivalence (1) is preserved, as is the accuracy of Russell's translation of " $\imath Fx$ exists" by "$(\exists y)(x)(Fx \equiv x = y)$" which in turn is equivalent to "$(\exists y)(y = \imath Fx)$". And since the evaluation rules require us to look only at the evaluation of the immediate constituent term and not at whether it is a variable, individual constant, or definite description, the language is rendered logically perfect.

It may be objected that all we have done is to find two ways of coding information about the syntactical character of a term (whether or not it is a definite description) into the evaluation rules. Thus if a term has no denotation or none within the domain of discourse, we know that it must be a definite description. But such an objection would be both inaccurate and short-sighted. Inaccurate, because it ignores the fact that the treatment of improper descriptions no longer distorts the treatment of proper descriptions, which now receive the natural semantical analysis. And shortsighted, because it ignores the possibility of using the semantical ideas of the modified Frege-Strawson method or the modified Frege-Carnap method to construct theories which depart still further from that of Russell but which allow a complete assimilation of definite descriptions and individual constants, and even provide for true atomic sentences about non-existing individuals. Indeed, such theories have already been constructed by a number of authors, among them Hintikka, and Lambert, and Scott [8].

Russell and Frege were both interested in removing the logical imperfections of ordinary language, but their methods were quiet different. (For what follows it is necessary to assume the 'translation-rule' interpretation of contextual definitions, according to which definite descriptions do not occur at all in the perfect language.) Where grammar called for entities whose nature was obscure, Frege attempted constructions, as with numbers, or a theory about the purported entities, as with propositions. Thus he sought to preserve the integrity of ordinary language by ontological ingenuity. Russell's response, at least in the case of definite descriptions, was by grammatical reconstrual and replacement. The two methods are easily contrasted by their analyses of "the number of planets is two" (I simplify for economy). Following Frege's method, we accept the apparent grammatical form (that of an identity sentence) and translate into something like "the similarity class of the set of planets = the set of all couples." Following Russell's method (though he might not follow it in this case),we dispense altogether with singular terms purporting to refer to numbers and translate into something like "there is a planet, and there is another, and there are no others." Either method may lead to logical perfection, but Frege's way seems to me more fruitful in the long run. I see Frege's method at work in Tarski's reduction of possible worlds to models,[12] Carnap's reduction of propositions to classes of models,[13] Wiener's reduction of ordered couples to classes [9] (and hence of relations to classes), and of course Frege's own treatment of numbers [10]. I would classify as applications of Russell's method: Stevenson's emotive analysis of "x is good" [11], Austin's reconstrual of the singular term "the-meaning-of-(the-word)-'rat' "[12], Quine's treatment of virtual classes [13], and, classically, Russell's own treatment of definite descriptions. The scope and limitation of Russell's method, in one of its applications, has received careful and extended discussion in Quine's *Set Theory and its Logic* [13].

Russell seems not to have viewed his linguistic replacements as being in sharp contrast with Frege's ontological constructions, for in his essay "Logical Atomism," he writes:

> One very important heuristic maxim which Dr. Whitehead and I found, by experience, to be applicable in mathematical logic, and have since applied in various other fields, is a form of Ockham's razor ... The Principle may be stated in the form: Wherever possible, substitute constructions out of known entities for inferences to unknown entities.
>
> The uses of this principle are very various, but are not intelligible in detail to those who do not know mathematical logic. . . .
>
> A very important example of the principle is Frege's definition of the cardinal number of a given set of terms as the class of all sets that are "similar" to the given set. . . . Thus a cardinal number is the class of all those classes which are similar to a given class. This definition leaves unchanged the truth-values of all propositions in which cardinal numbers occur, and avoids the inference to a set of entities called "cardinal numbers," which were never needed except for the purpose of making arithmetic intelligible, and are now no longer needed for that purpose. . . .
>
> Another important example concerns what I call "definite descriptions," i.e. such phrases as "the even prime," "the present King of England," "the present King of France." There has always been a difficulty in interpreting such propositions as "the present King of France does not exist." The difficulty arose through supposing that "the present King of France" is the subject of this proposition. . . . The fact is that, when the

[12]Implicit in "Über den Begriff der logischen Folgerung" [6].
[13]Reference [3], §40.

> words "the so-and-so" occur in a proposition, there is no corresponding single constituent of the proposition, and when the proposition is fully analyzed, the words "the so-and-so" have disappeared." [14]

In so far as we find reconstruction of our grammatical intuitions and reconstruction of our ontological intuitions equally congenial, this conflation of ontological construction with grammatical reconstrual is harmless. But in so far as our grammatical preconceptions continue to dominate our ideas, logical perfection achieved in Russell's way will remain unsatisfactory.

REFERENCES

1. B. Russell, "On Denoting," *Mind,* Vol. 14 (1905); reprinted in B. Russell, *Logic and Knowledge*, London, 1958, and also in *Readings in Philosophical Analysis* (H. Feigl and W. Sellars, eds.), New York, 1949.
2. G. Frege, "Über Sinn und Bedeutung," *Zeitschrift für Philosophie und philosophische Kritik,* Vol. 100 (1892); "On Sense and Reference" in G. Frege, *Philosophical Writings* (translated by Geach and Black), Oxford, 1952, and also as "On Sense and Nominatum" in *Readings in Philosophical Analysis.* See especially footnote 9.
3. R. Carnap, *Meaning and Necessity,* Chicago, 1946, especially §8.
4. P. F. Strawson, *Introduction to Logical Theory,* London, 1952.
5. B. Russell and A. N. Whitehead, *Principia Mathematica,* Cambridge, 1910.
6. A. Tarski, "Der Wahreheitsbegriff in den formalisierten Sprachen," *Studia Philosophica,* Vol. 1, 1936; and "Über den Begriff der logischen Folgerung," *Actes du Congrès International de Philosophie Scientifique,* Vol. 7, 1936. Translated as "The Concept of Truth in Formalized Languages," and "On the Concept of Logical Consequence," in A. Tarski, *Logic, Semantics, Metamathematics,* Oxford, 1956.
7. D. Kaplan, *Foundations of Intensional Logic* (Dissertation), University Microfilms, Ann Arbor, 1964.
8. J. Hintikka, "Studies in the Logic of Existence and Necessity: I. Existence," *The Monist,* Vol. 50 (1966); D. Lambert: "Notes on E!III: A Theory of Descriptions, *Philosophical Studies*, Vol. 13, 1962; K. Scott: "Existence and Description in Formal Logic," in *Bertrand Russell: Philosopher of the Century* (R. Schoenman, ed.), London, 1967.
9. P. P. Wiener, "Simplification of the Logic of Relations," *Proceedings of the Cambridge Philosophical Society*, Vol. 17, 1912-1914.
10. G. Frege, *Grundlagen der Arithmetik,* Breslau, 1884; translated as *The Foundations of Arithmetic,* Oxford, 1950.
11. C. L. Stevenson, "The Emotive Meaning of Ethical Terms," *Mind,* Vol. 46 (1937); reprinted in *Logical Positivism* (A. J. Ayer, ed.), Glencoe, Illinois, 1959.
13. W. V. O. Quine, *Set Theory and its Logic*, Cambridge, Mass., 1963.
14. B. Russell, "Logical Atomism" in *Contemporary British Philosophy* (J. H. Muirhead, ed.), London, 1924; reprinted in *Logical Positivism* (A. J. Ayer, ed.), Glencoe, Illinois, 1959.
15. B. Russell, *Introduction to Mathematical Philosophy,* London, 1919, Chapter XVI.

DISCUSSION

QUINE: (To Kaplan:) Concerning your notion of the perfect language, and the way in which the immediate context is determined by the immediate constituent: Is this determination effective?

KAPLAN: Consideration of the computability of the evaluation rules doesn't enter into my notion of logical perfection.

QUINE: It is not part of the definition?

KAPLAN: It is not part of the definition, although it is clear that in certain interesting cases you might want to impose that further condition, just as you might want to impose effective computability for the functions involved in building up the grammar, although in a very general case we do not.

YOURGRAU: I should like to raise a question in regard to the so-called logically perfect language. It is of course an artifact, because the language we talk is by definition not perfect. We therefore strive to develop an artifact.

Now, I learned from Carnap that if you have an idealized, formalized language, provided you have correspondence or transformation rules, you can express every situation, event, concept, etc. in an arbitrary language, colloquial or otherwise.

I would like to know whether you share the enthusiasm for, and the faith in, such a language which Carnap advocated so strongly. I understand this as an ideal goal, one that we would like to attain as logicians, but I am not convinced it can ever be be realized.

Twice in your paper you used the term "contextual definition," which is one of Church's favorite expressions. Do you mean to say that you can do away with the explicit definitions which the mathematicians desire and without which we have no physics? Are you satisfied with the contextual definition as long as we play the game according to the rules which you have just set out here? In other words, do you say, "I don't want to give any explicit, 'definite' definition?"

My last question is: Would you maintain that the defect in Russell's theory of descriptions, which had (with some justification) acquired the status of a scripture, is that it does not allow application to indefinite descriptions, although they are entirely legitimate and play such a large role in ordinary discourse? I mean not only in colloquial language but in specialized languages too, for instance in science. Would you therefore like to enlarge the range of Russell's theory of descriptions either by adding or patching up or even by eliminating whole sections of that "sacred text" of the logician?

KAPLAN: My comparison of Russell's theory of definite and indefinite descriptions was intended to be a *reductio ad absurdum.* It is quite clear that it would serve no useful purpose to introduce into the language an expression which grammatically functioned as a term but which had the meaning of "a girl," "a man," and so on, a term which was an indefinite description. I would not favor that.

On contextual definition versus explicit definition, I am all in favor of explicit definitions. Part of the point of my paper, part of the question that I was asking in the title, was an attempt to understand what a contextual definition really is. I suggested that contextual definitions don't satisfy one of the standard criteria for explicit definitions, that is, assigning meaning to the expression that is being introduced. Further, I suggested some other ways that we might think about these things. I think the phrase "contextual definition" is an extremely unfortunate choice for the kind of thing that is being done there. I prefer to think of the equivalences that are given in (1) and (9) in either the way that I suggested, as intermediate steps in translation, to help you to get the English under more control before you put it into the symbolic language, or else not to think of the description operators as being defined at all but to simply let this be an equivalence, an axiom. The equivalence would simply be an axiom within the language, and one would then try to find a way of doing the semantics.

Can everything be expressed in a logically perfect language? I don't see why not. This question suggests that the search for logically perfect languages—I have tried to make it clear that the notion of logically perfect languages is not a really terribly metaphysical notion—that the search for a logically perfect language is somehow an abandoning of the full expressive resources of the languages of everyday life. I don't think it is anything like that at all. The people who were after logically perfect languages were simply after languages whose syntax more clearly reflected what they were asserting. When we create a mathematical language, as can be seen quite clearly from some of my bilingual formulations here such as (11) and (12), there is an enormous increase in efficiency that is gained by going to symbolic languages. It is not that we can say something that was impossible to say before, but that we can say it more succinctly and more clearly, and the point about the logically perfect language isn't to make it shorter in any way, it is that it should make it clearer by having the grammar of the sentence somehow reflect more closely the meaning of the sentence.

You asked generally about Russell's theory of descriptions. I think a useful distinction that might be made is this one. Sometimes I distinguish what I call Russell's *theory* of descriptions from Russell's *theorems* of descriptions. Russell's theorems of descriptions involve a certain body of theorems formulated in a certain language. Russell's theory of descriptions involves a certain semantical analysis of that language. I think we can preserve Russell's theorems of descriptions without maintaining Russell's theory of descriptions.

TENNESSEN: I am always puzzled when we speak as though it was taken for granted that, for instance, Kaplan's formula (8) could be interpreted in the direction of P and $\sim P$.

It seems to me to be, at the face of it, a very unreasonable and implausible interpretation. It is clearly what I call an "analytical-form" sentence. Analytical-form sentences are never interpreted to be analytically either true or false. "A bachelor who isn't married" may still be interpreted as analytically true, but "a bachelor is a bachelor" can never plausibly be interpreted in that direction. It is clearly misleading (to say the least) to simply state: "(8) is a contradiction." What Kaplan probably intends to express is something like: "(8) is here meant to be interpreted as though it were an example of a (logical) contradiction." In other words, what he is really saying is: In spite of the obvious implausibility, let (8) be interpreted in such a direction that, if symbolized, it could be conveyed by $(P \cdot \sim P)$. Or is to be taken for granted that this is *the* interpretation of it?

KAPLAN: I agree that when we see somebody seriously asserting a sentence which is either trivially true or trivially false, the only plausible thing is to assume that he doesn't understand it in such a way that it *is* trivially true or trivially false. And so, looking at (8) (which happens to be a strikingly good example of just this kind of thing), I can see that the interpretation that Tennessen wanted to put on it is not this interpretation. So I agree on that.

However, I think the point I wanted to make with (7) and (8) could equally well have been made not by using the conjunction in the case of (8), which is in fact trivially false, but by describing a dialogue. I am trying to show that (5) and (6) are not parallel in their structure. And so instead of (7), I might have said: Imagine John saying, "A senator from New York is supporting Rockefeller," and Fred getting up and saying, "A senator from New York is not supporting Rockefeller," and, of course, they may be in agreement. On the other hand, imagine John getting up and saying, "Jacob Javits is supporting Rockefeller," and Fred getting up and saying, "Jacob Javits is not supporting Rockefeller." There the contradictory kind of interpretation is more plausible. I just wanted to point out that difference.

QUINE: Transformation and definition are related in certain ways, as Kaplan pointed out. If we bring in this transformation we no longer have what he calls the logically perfect system. However, we do have something that will fit into the pattern of transformational grammar. Also, if we could regard the transformation as a definition, we would have this relationship: The system would not be, as it stands, a logically perfect system, but it would be a logically perfect system plus a definition. Or, to put it the other way around, the definition would be instruction as to how to translate this non-logically perfect, this logically imperfect system into another which is logically perfect.

Then the question arises, whether in all transformational systems, the transformations could be regarded similarly as definitions, and therefore all transformational systems could be regarded as simply translatable into logically perfect systems. If the answer to that were "yes," I think people like Noam Chomsky would feel rather frustrated.

Then I am worried about contextual definition. We have to distinguish between contextual definition and implicit definition. Now, contextual definitions are explicit definitions as opposed to implicit ones, and, in fact, just about every definition that we ordinarily think of as an explicit definition is a contextual definition. We define "*P* or *Q*," for instance, as "not (not *P* and not *Q*)." I am not equating "or" to any expression. That would be an example of a definition which was not contextual. Or define "5" as "4 plus 1." This is not contextual either.

Here I am not giving anything that the "or" is equivalent to. I am explaining a context, and between this and Russell's treatment of descriptions I see no difference except of degree. Contextual definitions generally are explicit in the sense that they tell us exactly how we can translate a given notation into another notation that is free of the sign in question, whereas implicit definitions, so miscalled, aren't definitions at all; they are axioms, theorems, assertions of the kind that don't enable you, in the general case, to eliminate the signs or to translate into another language that is free of them.

KAPLAN: Concerning the question about the definition of "or," Quine claims that "(P or Q) $\equiv$ not (not P and not Q)" is also a contextual definition. That is, we have on the left of the equivalence, or whatever it is that forms the definition, not just the sign being defined but also some other expressions, and he said that this is a contextual definition.

I don't believe that the difference between this and (1) is simply a matter of degree. I contend that there is a radical difference. In the case of "or," the definition could be thought of as really producing an explicit definition. In my view the crucial thing is something like this: As long as we introduce just the new sign plus variables for the argument expressions, it can be transformed into an explicit definition of just the new sign, provided that the logic which is available in our language is strong enough. If we have the lambda operator, for example, you can take the variables over on the other side, and really just define the connective. I don't think that anything like that can be done for Russell's iota without changing its grammar from a term-maker to a formula-maker, because the variables on the left in (1) include "*F*," which is not an argument expression of iota. Of course, the whole problem of defining variable-binding operators is a little unclear anyway. But I think that in any adequate treatment of the problem, there will be an important difference between, say, (1) and the definition of the existential quantifier, and I think that the difference is connected with how the grammar changes when the 'contextual' definition is transformed into an explicit definition.

Now, Quine raises a very interesting point about changing the transformational operations in the semantics into the kind used in a logically perfect language. Probably it can be done for things like the iota and things like the biconditional, but I don't think it could be done in general. The method that I suggested for getting a logically perfect language, which would handle Russell's equivalence, seems to me not to be completely obvious (although a number of other people have thought of it, including, I believe, Hintikka and Scott).

I agree very strongly with Quine's point about implicit definition. One of the points I want to make is that the word "definition" gets thrown around carelessly. Maybe I could make the point in a different way. We should carefully separate different kinds of definitions because they have different logical properties which are of some importance. The so-called implicit definitions aren't definitions in any sense. We have a number of primitive constants, write down axioms on them, and people say, "Well, you have implicitly defined them." You have, if the system has only one possible interpretation. If it is categorical, then there is a certain sense in which one might say that they have been implicitly defined. But in general one hasn't defined them in any sense—they have just been left open; one has merely said a few things about them. If somebody asks me to define the description, "The present queen of England," I don't define that expression implicitly or in any other way by saying, "She is more than 32 years old."

HINTIKKA: I would like to register a protest, not so much against the theory of descriptions in any of its forms, as against some of the uses that philosophers have made of it. I think that the original form of Russell's so-called theory is a very nice illustration of the philosophical principles that Kaplan quoted, and therefore, in a way, a very nice example of what can be done. But I think philosophers, especially philosophers of language, have pushed it into roles where it doesn't always help and sometimes confuses the issues.

What I mean is roughly this: In analyzing, not the kind of context Russell was in the first place interested in, but what are known as modal contexts, there is a great temptation to think that we can get rid of some of the problems by saying, "Well, let's just construe the free singular terms we have to deal with as definite descriptions. Definite descriptions we can get rid of in terms of quantifiers, so that problems concerning singular terms other than bound variables get solved in this way."

What we get in this way often doesn't help us to understand the underlying problems at all. Take, for instance, the problem of seeing how singular terms behave in the context of modal operators. It may be that by some superhuman ingenuity one could perhaps handle this by replacing other free singular terms by definite descriptions. However, as a matter of fact, I don't think we can see through the situation without bringing in, at least for heuristic purposes, free singular terms and without forgetting the theory of definite descriptions for the time being. On the formal level some of the difficulties with definite descriptions arise by way of scope conventions in modal contexts. But Russell's theory itself does not explain how these scope conventions—which really are not conventional at all—are to be chosen, or why they are chosen differently on different occasions. There are even good reasons for assuming that no scope conventions can do full justice to ordinary usage. Russell's theory of definite descriptions has been called, justly, a paradigm of philosophy. I am afraid that some of the uses that have been made of this "theory" are nevertheless paradigms of bad theory and bad philosophy.

LEJEWSKI: At one stage in his paper Kaplan mentions an interesting sort of tension within the Russellian theory of descriptions. He reminds us that Russell has a symbolic notation for definite descriptions, a notation which he is anxious to introduce into the language of *Principia Mathematica*. In fact, he introduces it by means of contextual definitions. But contextual definitions, and definitions in general, are treated in *PM* as typographical conveniences, which are outside the system altogether. Thus it appears that the descriptional notation is made use of in *PM de facto* although *de jure* it is not to be there at all.

One of the fundamental presuppositions of the language of *PM* is that at its lowest categorial level it has proper names. Improper names are not allowed as meaningful substituends for the lowest level variables. Improper names have to be analyzed away, and the same applies to improper descriptions. There does not seem to be any theoretical necessity for analyzing away proper descriptions, but, as Kaplan points out, Russell preferred to treat them on a par with improper descriptions. If we agree to do without improper names and improper descriptions, can we still legitimately introduce proper descriptions into a language such as that of *PM*? The answer seems to be "yes," but instead of "abbreviational" definitions we have to make use of definitions which, as regards their status within the system, are more like axioms.

A system comparable to that of *PM* can be constructed in terms of a language which presupposes that proper and improper names as well as proper and improper descriptions are all acceptable as meaningful substituends for the lowest level variables. The system is based on the following single axiom, $[a\ b]\colon a=b\cdot\equiv\cdot[\exists c]\cdot c=a\cdot c=b$, and it has the following rule of definition:

On the assumption that a thesis T is, at the given stage, the last thesis of the system, we can add to the system a new thesis of the form

$$D.[a\ldots]\ \colon\!\colon a=\alpha(v_1,\ldots,v_n)\cdot\equiv\ .\!\cdot\!.\ \phi(a,v_1,\ldots,v_n)\ .\!\cdot\!.\ [b]\colon\phi(b,v_1,\ldots,v_n)\cdot\supset\cdot a=b\ ,$$

provided the following conditions are fulfilled: (1) α is a new constant functor; (2) $a, b, v_1, \ldots, v_n$ are distinct variables; in $a = \alpha(v_1, \ldots, v_n)$ each of $a, v_1, \ldots, v_n$ occurs only once; (3) $\phi(a, v_1, \ldots, v_n)$ has no free variables other than $a, v_1, \ldots, v_n$; (4) $\phi(b, v_1, \ldots, v_n)$ is equiform with $\phi(a, v_1, \ldots, v_n)$ except that it has occurrences of b instead of occurrences of a; (5) $\phi(a, v_1, \ldots, v_n)$ is, with respect to T, a meaningful expression of the system.

In formulating this rule I have omitted, for the sake of simplicity, a few minor refinements.

Now, it seems to be obvious that the Russellian $\imath x F x$ is a function of F, which in turn is its argument. If we use $\imath$ as the functor, then the function can be symbolised as $\imath(F)$, and the functor $\imath$, which forms a definite description in concatenation with F as the argument, can now be defined as follows:

$$D1.\ [a\ F] :: a = \imath(F) \cdot \equiv .\therefore F(a) .\therefore [b] : F(b) \cdot \supset \cdot a = b \cdot$$

Our definition is not a convention for abbreviating formulae. It introduces, into the system, definite descriptions *de jure*. It is to be noted that the problem of scope or the problem of primary or secondary occurrences of descriptions does not arise here.

Can our rule of definition be adapted to the language of *PM*? It seems to me that it can. We can, for instance, adopt the procedure whereby we are entitled to add to the system of *PM*, as a new thesis, the left-hand side of an equivalence satisfying schema *D*, provided the right-hand side of the equivalence has been proved to be a thesis of the system. This is only an outline of the procedure, whose details remain to be worked out.

Schema *D* shows how closely the problem of definite descriptions (and that of descriptive functions) is connected with the problem of definitions construed as means of introducing new vocabulary into a system.

KAPLAN: It seems to me worthwhile to investigate the consequences of admitting "definitions" of the form suggested into the system of *PM* and also into other kinds of systems.

I want to try to emphasize a certain aspect of my paper, though. Especially of late there has been a great interest in what is called semantics, that is, the connections between the syntax of a language and the interpretation of a language, and my paper was really on that topic. In particular, the notion of a logically perfect language seems to me to be a semantical notion concerning just this kind of relationship. The question of what is a definition and how different definitions can be introduced: contextual, implicit, abbreviation, and so on—these are all very interesting problems at the level of syntax. But I am mainly interested in the relation between writing out these different syntactical ways of doing it, the formulas, and the semantical interpretation that is put upon them.

QUINE: Concerning Hintikka's comment, I would just say this regarding the eliminability of definite descriptions and other singular terms (apart from variables themselves), which I insist on frequently: I am all for bringing them back in for heuristic purposes where they help. In so far as you can define something away, you also excuse its restoration. To define is to eliminate; it also is to exculpate.

KAPLAN: I think I am slightly on Hintikka's side on this matter and I believe that some of my remarks have something to do with this situation. Especially the idea of trying to develop the semantics so as to treat the language as being logically perfect—I want to leave out the whole business about transformational grammars for the moment. The idea is connected with a certain kind of integrity, taking the grammar of the language, the syntax, to be as it appears to be. Now, if we follow Russell's method, especially as developed by Quine—Quine has given us, in fact, the natural extension of it to take care of individual constants and has made the theory much more pleasing to the logicians—then the sentences that appear to have subject-predicate form, turn out on their analysis (whether it is by contextual definition or whatever, looking just at the semantics) not to have that logical form at all. Thus, so long as the original sentence remains in the language, it is not logically perfect. It is not that one cannot take a language without definite descriptions and get very interesting results for it, or that one cannot show that this kind of Russell-Quine elimination can be consistently performed. It is just the feeling that when that kind of transformation and elimination *is* performed, it is not absolutely clear that we really are talking about the sentences understood in the same way as they were in the first place. I suppose this applies, at least in part, to Russell's claim that the equivalence that he writes down, the so-called contextual definition, really does reproduce the sense or the meaning of the sentence that he is analyzing.

COHEN: I don't understand, as a non-logician, how this very important problem about equivalence is settled among logicians or among philosophers of the sciences who may offer logical reconstructions. When you disagree about whether what is to be defined and the proposed definition have in fact the same meaning, how does one settle this matter in logic? What is the methodology of settling a disagreement on whether the definition really does carry out what is claimed?

KAPLAN: I don't know that there is some way of settling these questions. Perhaps the thing to do is to develop the system, as Quine has developed it, and also to attempt to develop the system in which the language comes out to be logically perfect, or in some other way. Of course, you have to learn some logic so that you can understand what these developments are. You then look at the different developments, and if you want to, you can go ahead and be pleased more by one than by the other, or just be interested in both of them and choose the one that seems most appropriate for the problems that you have at hand.

It is always very interesting to discover that you may develop either of these systems in what seems to you to be an intuitively different way; and possibly at the end you can prove some interesting kind of equivalence between the two, or possibly not. But I don't think the question as to which of these really gives the correct sense of the word in the ordinary language is an interesting question until we have brought in things like the notion of a logically perfect language and said, "Well, this makes it logically perfect, this makes it contextual, this makes it so-and-so." Then we have some rigorous statements we can make, e.g. this definition has this property, it doesn't have that or another.