It is commonly supposed that Russell adopted the Theory of Descriptions in preference to Frege’s Theory of Sense and Reference entirely because of puzzles about the truth value of propositions like “The King of France is bald”. In fact Russell’s objections to Frege’s theory were much more extreme. In an obscurely written passage in “On Denoting”¹ he argues that the theory is an “inextricable tangle” and “wrongly conceived”. I shall try to reproduce Russell’s arguments in a way which will preserve their strength while discarding their unclarity and then I shall criticize them.

Russell’s argument proceeds from two explicit assumptions:

(1) When we wish to refer to the sense of a referring expression we do so by enclosing the expression in inverted commas.²

(2) The sense of a referring expression refers to the referent.

Russell takes (2) to be Frege’s thesis on the relation of sense and reference, and it is this which he tries to reduce to absurdity. An immediate consequence of it for Russell is the corollary

(2a) Whenever the sense of an expression occurs in a proposition, the proposition refers to the referent of that sense.

The reductio ad absurdum which Russell attempts to draw from Frege’s theory takes the form of a dilemma for the theory: either the relation between sense and reference is a logical relation, as expressed by (2), or it is not a logical relation. But in the former case it becomes impossible to refer to the sense of a referring expression, and in the latter case, though we might claim to be able to refer to the sense of a referring expression, the relation between the sense and reference of an expression remains


² Russell uses the expressions ‘meaning’ and ‘denotation’ respectively for Frege’s ‘Sinn’ and ‘Bedeutung’. I have adopted the more modern translation ‘sense’ and ‘reference’ or ‘referent’ and substituted them for Russell’s terms throughout. Cf.TRANSLATIONS FROM THE PHILOSOPHICAL WRITINGS OF GOTTLIB FREGE, trans. Geach and Black, pp. 56 ff.
“wholly mysterious” and the theory has no explanatory value. Thus if the theory is explanatory it breaks down and if not explanatory it is philosophically worthless.

The crucial proposition then which Russell must prove is that on the assumption that (2) is true it becomes impossible to refer to the sense of a referring expression. He proves this by means of examples, and (1) is used as an aid in presenting the examples. As such an aid however it suffers from two serious defects; first it is false and therefore its bald assertion is confusing to the reader and produces confusion in what follows. A systematic account of the use of inverted commas is difficult to give because of the wide variety of uses to which they are put in ordinary speech, but it does not seem to me that there are any contexts at all in ordinary speech where enclosing an expression in inverted commas is by itself sufficient to indicate that the resultant expression is being used to refer to its customary sense; rather, as Frege said, do we use the expression “the sense of the expression ‘.......’” to refer to the sense of an expression. And secondly even if (1) were true, or were taken by the reader to be true, or were treated by the reader as an arbitrary ruling introduced by Russell to aid him in presenting his argument, its subsequent employment is confusing, for Russell uses inverted commas in at least three different ways: as indicating that expressions are being presented and spoken of rather than used in their customary way; as “scare quotes”; and as a means of referring to the sense of expressions.1 Furthermore on some occasions Russell does not use quotes where the context indicates that they should be included. Thus in order to give a clearer exegesis of Russell’s argument I shall adopt a variation of (1) as an arbitrary convention to be employed for purposes of the argument. I shall rewrite (1) to read

(1) Whenever an expression occurs surrounded by the squiggle signs thus § ...... §, the resultant expression is to be taken as referring to the sense of the original expression. In other words, for any expression, say “the dog”, the sense of the expression “the dog” = § the dog §.

We are now in a position to restate Russell’s arguments for the conclusion that on Frege’s theory it becomes impossible to

1 Church supposes that Russell’s carelessness with quotes is sufficient to vitiate his argument; see his review of Carnap, Philosophical Review, 1943 (Vol. LII) p. 302. Russell’s argument however can be stated independently of these ambiguities, as was first suggested to me by Miss G. E. M. Anscombe and Mr. M. E. Dummett.
refer to the sense of a referring expression. Russell’s method here is to exhaust the possibilities by showing that each locution one could propose will fail of its purpose. Let us suppose that we wish to refer to the sense of some expression, e.g. “the dog”. Now if we wish to refer to its sense we cannot use the expression “The sense of the dog”, for to do so would be to refer to the sense of some dog, which being an animal and not a symbol, does not have a sense, and even if it did as part of some tableau vivant, that would not be the sense we intended. It seems we must use the expression “the sense of ‘the dog’” which by (1) is the same as “§ the dog §”. But, and this is the crucial point, these expressions won’t do either, for if the sense in question were genuinely referred to it would then become a constituent of the proposition, it would occur in the proposition, and by (2a) if the sense occurs in a proposition reference is then made to the referent of that sense, viz., the dog, and we have not succeeded in referring to the sense we wanted. In Russell’s words, “the moment we put a denoting complex [i.e. sense] in a proposition the proposition is about the denotation (of that complex)”.

Nor will it solve our problem to use any of the phrases, (a) “the reference of the dog”, (b) “the reference of ‘the dog’” or (c) “the reference of § the dog §” for (a) refers to nothing since the dog being an animal does not have a reference, (b) refers to the dog, and by (2a), so does (c) unless it is the same as (a) and has no reference. In any case we have exhausted all the possibilities without having succeeded in referring to what we wanted.

To put it briefly take any referring expression, say “X”, then its sense will be called “§ X §”, but paradoxically enough whenever “§ X §” occurs without inverted commas it must refer to X and not to the sense of the expression we wanted it to refer to.

Thus to speak of the sense, our subject must not be “§ X §” but something which refers to the sense which we wanted “§ X §” to refer to. Let us ad hoc introduce an expression to do this job, say “Y”. But now what is the relation between “Y” and its reference? “§ X §” cannot be a constituent of any definition we could give of “Y”, for we saw that whenever “§ X §” appeared without inverted commas, it was the reference of the sense and not the sense that was referred to. And of course we cannot argue backwards from the reference to the sense, since there are an indefinite number of senses attaching to any reference. Thus if “Y” really does succeed in referring
to a sense, the relation of sense and reference cannot be a logical one, for if it were, reference would drop straight through the sense to the reference of that sense without stopping. But if it is not logical, then the relation of "Y" to its reference remains wholly mysterious. Either Frege's house of cards collapses or we can only pretend to ourselves that it has not collapsed by succumbing to mysticism. Thus Russell's argument.

Two remarks should be made about where my exegesis differs from the original. First, the explanation in the text suffers from a pragmatic paradox in that Russell has to refer to what his argument claims cannot be referred to in order to specify what cannot be referred to. I try to avoid this by a freer use of the formal mode. Secondly, the original suffers from a slipshod use of the phrase "denoting complex". Russell introduces the phrase in a way which indicates that a denoting complex is identical with a sense, thus "denoting complex" just provides an alternative way of speaking, but he then goes on to speak as though a denoting complex bad a sense—a view which is inconsistent with his explanation of the phrase. I adopt his original view and ignore any phrases of the argument which depend on the other use as they seem to be nonsensical. Furthermore, no such phrase can be found in Frege: either a denoting complex is a sense, in which case the phrase is superfluous, or it is not a sense, in which case it is totally unclear and apparently irrelevant to Frege's argument.

Perhaps it will help us to understand the drift of Russell's argument if we can state it in more graphic terms. Imagine a game where marbles are dropped into bowls through pipes. This act is called referring. Pipes (senses) lead to bowls (references). It is a rule of the game that anything can be referred to. The difficulty is though that we cannot live up to this rule because we cannot refer to a pipe. Every time a marble drops into a pipe it goes through to the corresponding bowl. And it's no use saying we can construct a second pipe leading to the first for if the second pipe really leads into the first the marble will drop through to the bowl. And if it is not connected with the first, then how can it be used to refer to the first in accordance with the rules of the game? And of course we cannot get to the pipe by going to the bowl first for every bowl has an infinite number of pipes. 'There is no backward road from denotations [references] to meanings [senses]'. The way we have set up the rules of the game involves an absurdity. We must invent a new game (descriptions) that will eliminate referring.
Russell's Objections to Frege's Theory

Russell has a second argument emerging from the first. Not only, he says, does the reference of an expression occur in a proposition, but the sense as well. This is proved by example: "Scott is Scott" expresses a different proposition from "Scott is the author of Waverley", for the latter proposition possesses a property not possessed by the former, namely that George IV wanted to know if it was true. Yet on Frege's view (or rather Russell's version of Frege's view) only the reference of an expression occurs in a proposition. Hence on Frege's view the two propositions would have to be the same. But they are not the same, therefore Frege's view must be false since it entails a false proposition.

Russell's arguments suffer from unclarity and minor inconsistencies throughout and I have tried to restate them in a way which avoids these. But even in their restated form, they are faulty. Their faults spring from an initial mis-statement of Frege's position, combined with a persistent confusion between the notions of occurring as a part of a proposition (being a constituent of a proposition) and being referred to by a proposition. The combination of these two leads to what is in fact a denial of the very distinction Frege is trying to draw and it is only from this denial, not from the original thesis, that Russell's conclusions can be drawn. Let me explain.

Russell expresses Frege's principle by saying that the sense of a referring expression refers to the reference. Now strictly speaking this is not what Frege says. According to Frege,\(^1\) not the sense, but the sign refers and it refers in virtue of its sense, the sense provides the mode of presentation of the referent. A sign expresses its sense and refers to its referent. Russell's way of putting it might seem a pardonable metonymy on Frege, were it not for the fact that Russell adds to it the inexplicit assumption that if an object is referred to by a proposition then that object occurs as part of that proposition. The crucial stage of the argument, you recall, was that when "§ X §" is used in asserting a proposition then a sense is referred to by a proposition, but if a sense is referred to then it occurs in a proposition, but if it occurs in a proposition then its referent is referred to by the proposition (by (2a)), hence "§ X §" does not after all succeed in referring to a sense. Furthermore if its referent occurs as part of the proposition then it and only it occurs as part of the proposition, the sense being relegated to some higher limbo outside the proposition altogether. All this rests, I suggest, on

\(^1\) op. cit., pp. 57 and 61.
an equivocation between the notions of referring, and occurring in a proposition. Once these are clearly distinguished, the argument collapses. It is not easy at first sight to understand what might be meant by occurring in a proposition, except perhaps that words occur as part of a sentence used to assert a proposition. I think though that a plausible account can be given of this expression: it is equivalent to Frege’s notion of what is expressed in a proposition. To say that a sense occurs in a proposition is to say that that sense is expressed in the proposition. This explanation is of course far from clear—it is no clearer than is Frege’s explanation of the notion of expressing—but whatever its unclarity it is certainly different from the notion of referring. It is hard to find any conceivable sense in which a referent occurs as a part, or a constituent of a proposition. The tenth King of France is not a constituent of propositions about him the way his elbow is a constituent of his arm or he a constituent of a French wrestling team. Any view that makes the objects referred to by a proposition literally parts of that proposition is bound to be nonsensical. And even if we could assign some arbitrary rule that “occurring in” was to be treated as equivalent to “referred to”, this sense of occurring in is quite different from the one which is equivalent to “expressed by”, hence any arguments which rest on equating them are guilty of sheer equivocation. In other words, Russell’s argument which purports to develop Frege’s thesis in fact develops the negation of that thesis, for Russell’s assumption that occurring in a proposition is the same as being referred to by that proposition is an equation of sense and reference, and the whole point of Frege’s theory is to assert a distinction between sense and reference.

Once Frege’s intentions are kept clearly in mind, the puzzles about referring to the sense of an expression dissolve. The sense of any expression can be referred to by such a phrase as “the sense of the expression ‘E’” and the sense in question does not occur as part of the proposition in which this phrase is used, rather it is referred to by the phrase in virtue of the sense of the phrase, viz. the sense of the phrase “the sense of the expression ‘E’”. In short, the sense of an expression occurs in (to use Russell’s expression) a proposition, and in virtue of that sense the proposition refers to the referent. The referent does not occur in the proposition. Thus when a sense is referred to, the sense does not occur as part of the proposition, only the sense in virtue of which it is referred to occurs, thus the referent of the first sense is not referred to at all by the proposition.

Similarly Russell’s second conclusion, that Frege is unable
to account for the difference between “Scott is Scott” and “Scott is the Author of Waverley”, collapses. It only seemed a valid conclusion because of the equation of “occur in” and “refer to”. Once these are distinguished—and it is the whole point of Frege’s theory to distinguish them—it is easy to see how the two sentences express different propositions: they contain expressions with different senses.

Conclusion: Russell does not succeed in performing a *reductio ad absurdum* of Frege’s distinction but only of the conjunction of the distinction and its negation.

*Christ Church, Oxford.*