belief attributions that are appropriate or correct may change with a shift in the context of attribution. In the Ortcutt example, it is clear that, and how, the combining of the two halves of the story changes the context in relevant ways.

In general, what the pragmatic strategy suggests is a change in perspective. If we focus on the function of a belief attribution—to describe correctly and informatively a belief state by distinguishing between relevant alternative possibilities given by a context—then we can get a clearer picture of what is wrong with *de re* belief attributions in Ortcutt situations, and more generally, a clearer picture of the way we attribute beliefs and other representational mental states.

Comments

Hans Kamp

I.

As a rule the information carried by a linguistic expression depends on the context in which it is used. What a singular noun phrase denotes, or what proposition is expressed by a sentence, may vary from one context of use to another. But the ways in which expressions depend for their semantic values on the contexts in which they are uttered appear to be systematic. So it is one of the tasks of semantic theory to articulate how semantic value is determined by the interaction between syntactic form and contextual structure.

One kind of contextual dependency, which was investigated with much zest and some thoroughness at a number of California institutions in the late sixties, has to do with such contextual features as the time and place of the utterance, or the identity of speaker and addressee. These features determine the reference, in context, of the so-called "indexical" expressions—among them *I, you, here, now, yesterday* and the tenses of the verb. Essays that have come out of that particular tradition of research sometimes suggest that this is the only kind of context dependency one finds in natural language. Inasmuch as the suggestion is there, it is certainly misleading. For there is another kind that is as ubiquitous as the first, but which the Californians ignored. Stalnaker, I believe, has been the first to clearly recognize and articulate this other kind of context dependence. That recognition has been fundamental to a number of his earlier papers, and it is also the foundation for the proposals he makes in the present one. Indeed the present essay outlines the principles of linguistic meaning and context as Stalnaker sees them. Let me quote by way of illustration, just one brief passage:

"In the setting of the possible worlds framework, the kind of representation of context that we need . . . is provided by a set of possible worlds representing the *presumed common background knowledge of the participants in the conversation*" (my italics). It is clear from what Stalnaker says that this is to include "indexical" information, say concerning who is speaking, or when. But it is also clear from the surrounding text that it will usually include much more besides. In particular, it will often include information that has been put forward and accepted during earlier stages of the discourse of which the current utterance is part. Since the context may include information that earlier parts of the discourse have contributed to it, it will typically change as a conversation unfolds, or as a text reveals itself, sentence for sentence, to a reader. This dynamic aspect of contexts is crucial to an understanding of sentence interpretation generally and of certain aspects of belief attribution in particular.

Stalnaker's theory of contexts and their role in utterance interpretation is also an essential ingredient to his view of propositions. The concept of a proposition he advocates is not an absolute one, determined by criteria of identity that are fixed once and for all, but one that is flexible and context dependent. Any given context determines, among other things, a set of possibilities: the propositions relative to that context correspond to various ways of dividing that set up into two halves—the set of possibilities that are compatible with the proposition and the set of those that are not. An assertion u made in a given context C determines, by mechanisms that involve the syntactic form of the sentence uttered as well as various aspects of C, a proposition $P(u,C)$ relative to C; and the effect of accepting that assertion is to incorporate it into the context. As a result C will be modified into a new context C' in which the possibilities incompatible with $P(u,C)$ have been discarded. In other words $C' = P(u,C)$.

All this is both important and right. But there are, I think, some difficulties with the execution of the program that it suggests, i.e., that of developing an account of meaning for a natural language such as English along the indicated lines. These difficulties arise when contexts are identified with sets of possible worlds. This is an identification that Stalnaker seems willing to make. For instance, he says: "the kind of representation that we need to play these roles is provided by a *set of possible worlds* . . ." (my italics). Stalnaker has argued for this identification in other places, and he motivates it again in the present paper. His motives cannot easily be dismissed. Nevertheless I think that representations of the information on which language users rely and are meant to rely when they interpret what they hear or read must provide more detail than a set of possible worlds can deliver.

Here is an example, due to Barbara Partee, to illustrate the difficulty I have in mind. Compare the following two sentence pairs:

(1.a) Exactly one of the ten balls is not in the bag. It is under the sofa.

(1.b) Exactly nine of the ten balls are in the bag. It is under the sofa.

In (1.a) the pronoun *it* can be understood as referring to the missing ball. But in (1.b) this interpretation seems impossible. The only possible way we can understand this second *it* is to take it to refer to the bag (an interpretation which is also possible in the case of (1.a) although there it is—for reasons that do not matter here—not the one that first comes to mind). Thus the very same sentence—i.e., the second sentence of (1.a) and (1.b)—yields different interpretations in the contexts C_a and C_b induced by the first sentences (1.a.1) and (1.b.1) of (1.a) and (1.b).[1]

Let us suppose that C_a and C_b arise through the assertion and adoption of the sentences (1.a.1) and (1.b.1), respectively, in one and the same initial context C_0. Suppose also that C_0 is given as a set of possible worlds and that each of these sentences has the effect of cutting down this set to a new set consisting of just those worlds compatible with the proposition which that sentence determines in C_0. Since the sentence whose interpretation is at stake is the same in the two cases, the difference in interpretation must point to a difference between C_a and C_b. But if these contexts are sets of possible worlds, how could they be different? The two assertions that induce them, that of (1.a.1) and that of (1.b.1), are compatible with exactly the same possible worlds—those worlds in which nine of the ten balls are in the bag and the remaining one is not. So, if propositions are sets of possible worlds, the two assertions express the same proposition P. (Or, more accurately, they will, against the background of a normal initial context C_0, express the same proposition Pc_0.) So the resulting contexts C_a and C_b will be both equal to Pc_0, and thus will be equal to each other.

We must conclude that no difference can be predicted if contexts are identified with sets of possible worlds. Therefore a theory of meaning and context dependent interpretation of English must, if it is to handle such examples successfully, adopt a representation of contexts that goes beyond what sets of possible worlds are able to reveal.

Let me describe, albeit only in the barest outline, a theory that does not suffer from the difficulty we have noted. This is a theory that has been presented in formal detail elsewhere,[2] and here I will say just enough about it to make clear how it handles the problem which (1) presents, and to lay the basis for what I want to say below in relation to belief. The theory exists in two closely related forms, viz., that of the *File Change Semantics* of Heim and the (perhaps infelicitously named) *Discourse Representation Theory*, or *DRT* for short, which is presented in the remaining references mentioned in the last note. I will stick with the latter version, partly out of habit, and partly for reasons that cannot be explained at this point (cf. note 11).

In basic structure and assumptions DRT accords with Stalnaker's pro-

gram as I have sketched it. But, at least in the versions of it that exist in print at the present time, its aim has been somewhat more limited. The principal concern of these versions has been to provide the foundations of a theory of meaning and interpretation of discourses and texts, and moreover only of those which are made up entirely of indicative sentences, and can be treated as sequences of successive assertions. The theory assigns to each such sequence $<S_1,...,S_n>$ a set of truth conditions—or, if you accept possible worlds, a set of possible worlds, viz., the set consisting of those worlds in which the conditions are satisfied. This assignment is defined by a combination of two separate components. The first of these consists of a set of rules which, when applied to a given sentence S_i and a context structure K_i, produce a new context structure K_{i+1}. K_i is intuitively the representation of the context in which S_i is interpreted. K_{i+1} is the result of interpretation, a structure in which the information contributed by S_i has been incorporated into the contextual information identified by K_i, and which at the same time is the context of interpretation for the next sentence S_{i+1}. The structures K_i are called *Discourse Representation Structures*, or *DRS*'s. The rules and the principles according to which they are applied constitute the *DRS-construction algorithm*. The word "algorithm" is justified insofar as the interpretation procedure is indeed "computational": it operates according to explicitly stated canons on the combination of the representation structures K_i and the syntactic parses, according to some suitably chosen theory of syntax, of the sentences S_i. (Of course, we cannot expect the algorithm to be deterministic; for in actual fact many sentences and sentence sequences are ambiguous. We already noticed this in connection with the two sentence text (1.a). In application to such sequences one would not *want* the algorithm to be deterministic.) The second component is a truth definition, similar to those one finds in model-theoretic semantics. It applies, however, not to sentences or texts of the target language directly, but to the DRS's into which the construction algorithm converts them.

This much is in perfect agreement with Stalnaker's insights about context dependence. If there is a difference, it relates to the form of the DRS's. Each DRS determines a set of possible worlds; but it cannot in general be recovered from that set. This property of DRS's is forced upon them by the *double* role they play in the theory. They do not only determine the content (in particular the truth conditions) of what has been interpreted already, but must also serve as contexts for the interpretation of what follows. The construction algorithm must be able to operate in an essentially mechanical manner and this imposes non-trivial constraints on the structure a DRS can have. Indeed, the question of DRS structure, that of the form of the construction algorithm, and finally that of the syntactic structure of the sentences that together with DRS's provide the inputs to the algorithm, are all connected. Only a careful look at existing treatments of actual language
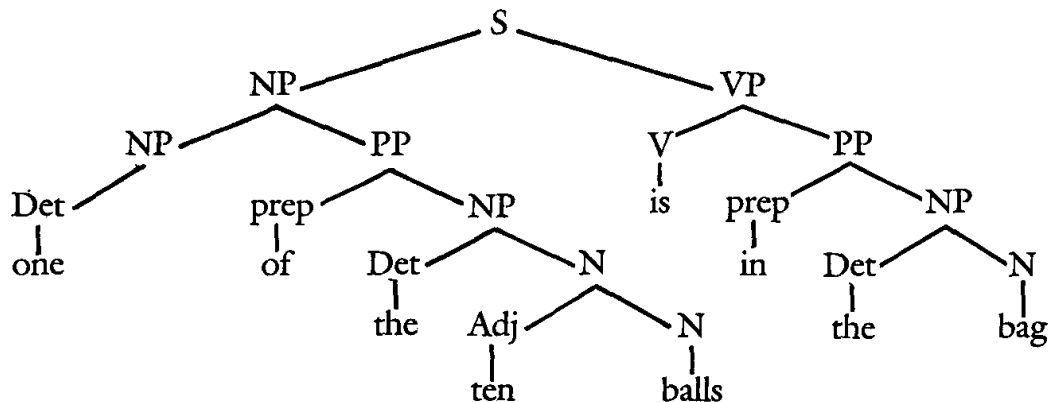
fragments will reveal how closely these problems, especially the first two of them, are interdependent.

A sketch of how DRT deals with the examples (1.a) and (1.b) will, I hope, give some idea of how the theory works generally; and in what sense DRS's can provide information over and above that which can be captured by sets of possible worlds.

First (1.a). Let us suppose that its first sentence is interpreted without any contextual background at all. (This is a situation which rarely if ever arises in actual discourse. But it is an assumption which simplifies the presentation, and the simplification affects none of the points I wish to illustrate.) Formally this means that the DRS K_0 which serves as context for the interpretation of this first sentence is the empty structure.

The construction algorithm can be stated in a number of different forms. In the version first articulated in my 1981 paper (see note 2), which I shall follow here, it operates on parses that have the form of trees and works its way down from the root, which represents the entire sentence, to the end nodes, which represent its minimal constituents. Let us make the largely uncontroversial assumption that the syntactic structure of the first sentence of (1.a) has (roughly) the following form:

(2)

```
                             S
               NP                        VP
        NP           PP              V         PP
   Det            prep    NP         is    prep     NP
   one             of   Det    N           in   Det      N
                        the  Adj    N            the     bag
                             ten   balls
```

The first few steps the algorithm performs when applied to (2) (and K_0) decompose the sentence into subject and verb phrase and analyze the complex subject term further. This leads, as in all other cases where a rule deals with one or more noun phrases, to the introduction of so-called *reference markers*. The roles these reference markers play will become apparent as we proceed. For the moment it is enough to note that they act as place holders for the noun phrases in the argument positions these occupy in the sentence, much—but as will be seen not altogether—like the variables of for-

mal logic. This means among other things that the step or steps which introduce the markers also must introduce certain items that identify which markers occupy which positions. These items are called *conditions*. Applying the first few steps to (2) we obtain a structure of the form:

(3)
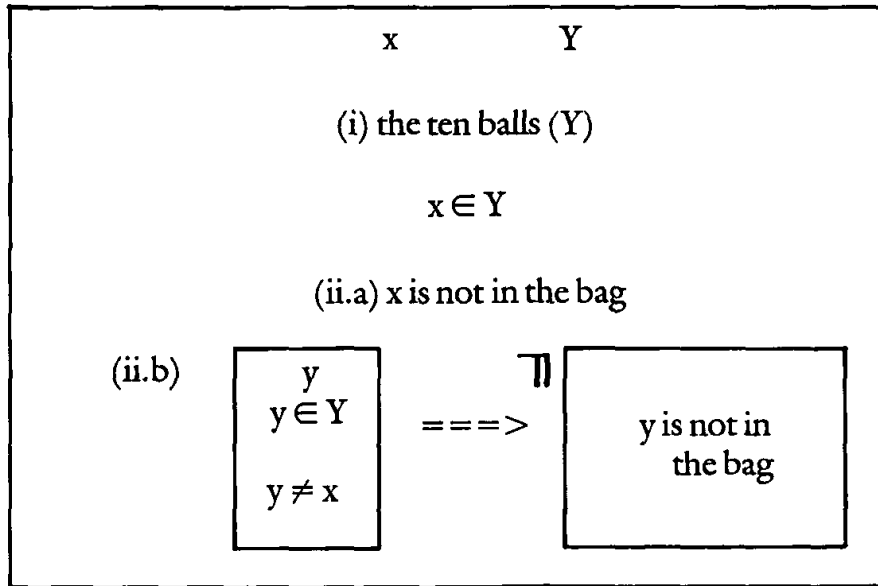
(The meaning of (ii.b) is that no ball other than x is not in the bag.)

This structure is not yet a completed DRS but an intermediate structure which the algorithm can convert into a DRS by breaking down the conditions (i) and (ii) further. However, as these further steps are not to the purpose of this illustration I will not discuss them.[3]

The truth conditions of (3) (and similarly for the complete DRS (4) that can be derived from it), are determined as follows: (3) is true, in a given world W, if it is possible to associate with the reference markers of (3)[4] corresponding objects of W which satisfy all the conditions of (3). It is in this, and *just* in this sense that the markers x and Y can be considered as "representing" objects. (The set of reference markers displayed at the top of the DRS, for which appropriate correlates must exist if the DRS is to qualify as true, is called its *universe*. The difference between the lower case x and the upper case Y is to capture the fact that x represents an individual and Y a set.)

The second sentence of (1.a) must now be interpreted in the context (3) provides. The first step in the application of the construction algorithm to the new sentence deals once again with its subject. This time the subject is the pronoun *it*. It has been said of pronouns that they must be identified

either with something' that is present in the environment in which the discourse takes place, and which is sufficiently salient (usually because the speaker points at the thing, or demonstrates it in some other way); or alternatively the pronoun must be identified with some element of the antecedent discourse or text. Often these uses are distinguished as the *deictic* and *anaphoric* use of pronouns. Though in some ways distinct they are closely connected, and in fact they seem to go together generally (we find them jointly associated not only with pronouns but also with definite descriptions and demonstratives; and this is true not just for English but for many other languages as well). Stalnaker's conception of context, in which both information about the environment and information deriving from the antecedent discourse are included, suggests why this might be so. This suggestion can be given a more tangible form within DRT if we make the reasonable assumption that both discourse information and environmental information are represented with the help of reference markers. The two uses can then be seen as variants of one single use, which always involves the resumption of a marker that the context-DRS already contains (or that, in certain cases which we will not encounter here, is being introduced into it simultaneously, but on independent grounds).

Of the two markers that the context DRS (3) supplies, only x is a suitable antecedent for the singular pronoun *it*, as the plural marker Y, which stands for a set of individuals, is excluded on categorial grounds. Thus we arrive at the interpretation which links *it* to the subject of the first sentence.[5] The result is the structure:[6]

(5)
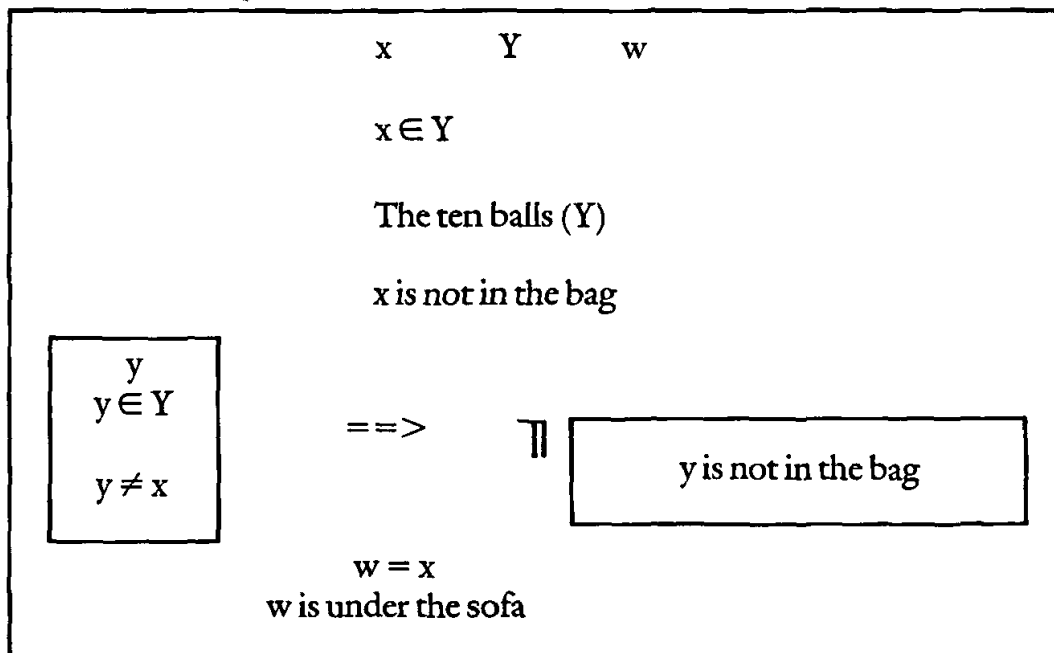
If we submit the first sentence of (1.b) to the construction algorithm, we obtain, after the first cluster of steps, the DRS

(6)
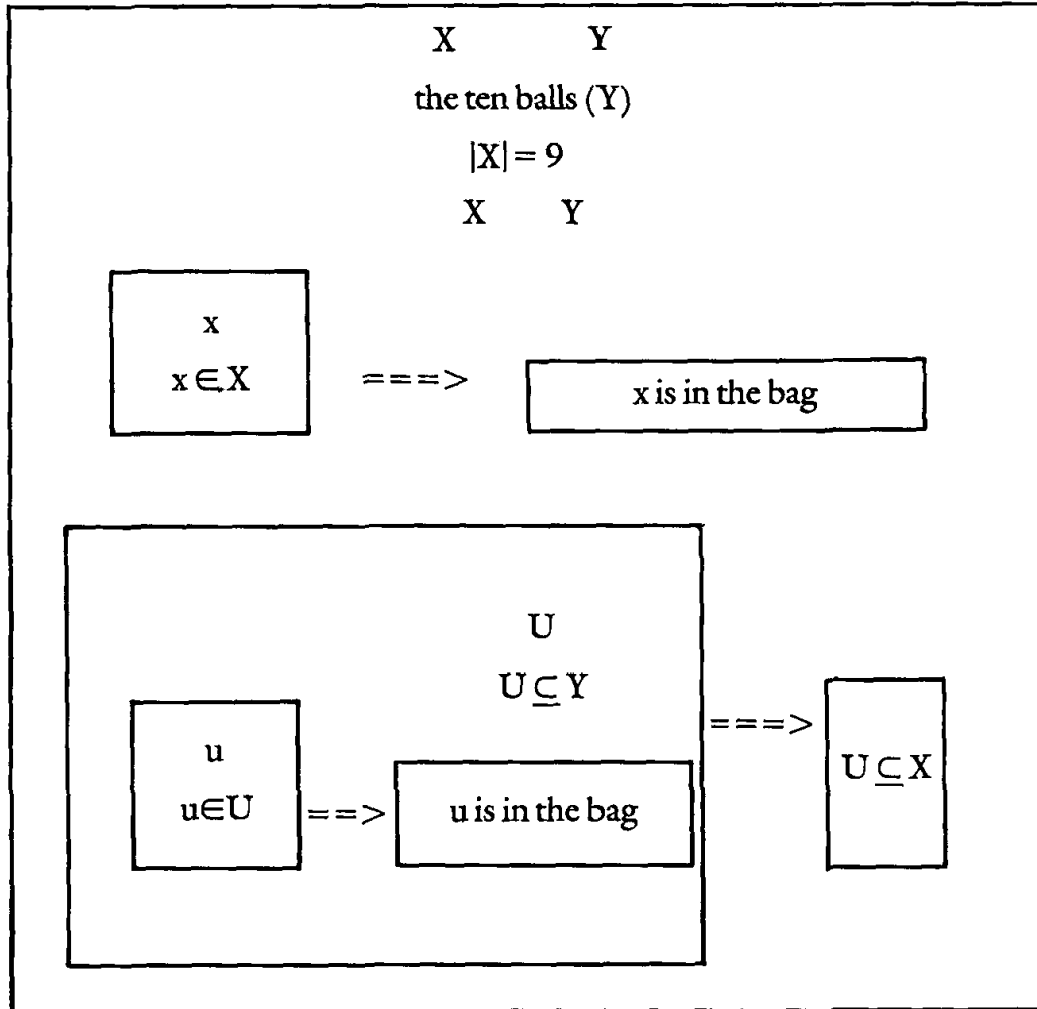
Once again, the precise form of the conditions here is immaterial for our purposes. The difference with.(3) that matters is that this time no reference marker for an individual ball is available. Consequently there is no way of interpreting the pronoun *it* as referring back to a ball. The only possible interpretation is that which takes *it* to refer to the bag, an interpretation the construction algorithm will in fact produce after it has completed (6) along the lines illustrated in note 3, and introduced a reference marker to.replace the NP *the bag*.

This is only a sketch. To mount a persuasive argument that the theory handles (1.a) and (1.b) correctly it would be necessary to look much more carefully at the construction rules that must be applied in the production

of the DRS's displayed in the diagrams in (3), (4), (5) and (6), and in fact at the entire construction algorithm that contains these rules as components. This is something which it would make no sense to carry through in these comments and so I must refer the skeptical reader once more to the cited works on DRT. To proceed further, however, we will have to take it as established that the problem (1.a) and (1.b) present can be solved along the lines I have indicated, and in particular that an adequate representation of contextual information must have at least the structural complexity of the DRS's we have so far encountered.

One further comment before we come to the subject of belief. Consider the DRS (5). The components of this structure stem partly from the first sentence of (1.a) and partly from the second. (The marker w and the conditions containing it come from the second; all the others from the first.) Thus (5) can be divided into two substructures, one consisting of the elements introduced by the first sentence and one consisting of those introduced by its successor. The first of these is just (3). The second one I specified in the parenthetical sentence above; its diagram would look like this:

(7)

$$
\boxed{
\begin{array}{c}
w \\[4pt]
w = x \\[4pt]
w \text{ is under the sofa}
\end{array}
}
$$
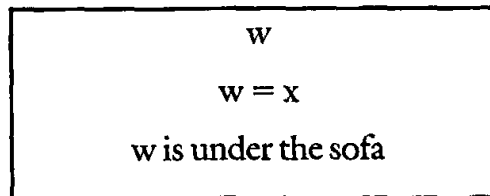
According to the truth definition of DRT which I outlined earlier, the DRS (3) seems to determine exactly the truth conditions, or if you like, the proposition, which (1.a.1) intuitively expresses. But what about (7)? Strictly speaking the truth definition cannot be applied to this structure. For suppose we associate w with the object b. What would it be for that object to satisfy the condition 'w = x'? Whether it does would evidently depend on what object is associated with x. But this is a matter that the truth definition does not provide for, as x does not appear in the universe of (7). We might adapt the definition so that it applies to structures such as (7) too. But how should it be adapted? If we change it by permitting it to treat x as if it were a member of the universe, the resulting truth conditions will clearly be at variance with what we take the second sentence to assert in the context of the discourse (1.a). For the truth conditions would now amount to no more than the claim that something or other is under the sofa, a claim which would be satisfied if, say, Fido was hiding under it, or the slipper you have

been looking for since yesterday, while the ball was in some entirely different place.

To cut a long story short, there is no way of adjusting the truth definition so that it handles DRS parts like (7), which consist of just those components of a larger DRS that come from a particular sentence of the discourse which the larger DRS represents, correctly. No matter how we modify the definition, the truth conditions it assigns to some such parts will misrepresent the content of the corresponding sentences. This indicates that we cannot hope to give a semantics of natural language discourse which treats each next sentence as expressing a proposition that identifies precisely the contribution which that sentence makes to the content of the discourse in which it occurs, without incorporating at the same time some or all of what was contributed by the preceding sentences. We cannot hope to develop an adequate theory which assigns such propositions even if we allow information from the antecedent discourse to guide the process of interpretation of the sentence to which it must assign a proposition. We cannot even develop such a theory if we do not equip it with representations for contexts that discriminate as finely as DRS's. No matter how much information about contextual structure the theory permits itself to take into account, it won't be able to always identify an absolute proposition that correctly captures just the contribution which the sentence makes to the content of the whole discourse or text. Our example makes this plain: All there is in the way of a proposition expressed by the second sentence of (1.a) is a proposition *relative* to the context induced by the first sentence and represented by (3).

So our example serves a second purpose as well as the one for which I brought it up in the first place. It illustrates Stalnaker's claim that one should in general see the semantic contribution a sentence makes in terms of a proposition that is a *subset* of the set of possible worlds determined by the context in which it is placed. This is an important observation, and one which plays a central part in the analysis Stalnaker proposes of belief attribution.

To conclude this section let me summarize its main points. I noted in connection with the "ball" example (1) that the contextual information which guides sentence interpretation is not always recoverable from the set of possible worlds the context determines. This seems to represent a certain difficulty for the way I took Stalnaker to conceive of the form of a systematic theory of meaning in context. And I sketched a theory which, by principles that are independently motivated, arrives at an account of contextual structure which enables it to explain the difference between (1.a) and (1.b), a difference which the possible worlds approach leaves a mystery. I then noted that the same example also serves to illustrate a quite different

point, which in fact is an important ingredient of Stalnaker's theory of the semantics and pragmatics of natural language: the need to replace the older absolute notion of proposition by that of a proposition *relative to* a context.

At this point, then, it looks as if our example vindicates Stalnaker on one point and refutes him on another. I want, however, to add a comment to the second, negative half of this double claim. There are various passages in Stalnaker's paper where he acknowledges that the representations of context which a comprehensive theory of meaning and use will have to employ must make explicit various bits of information that are not, in any obvious way, recoverable from a single set of possible worlds. For instance, he states[7] that the theory should explain, on the basis of the representation it offers of the context, what the *point* of a given utterance is. And he notes that such an explanation will need to refer to matters such as the presumed gap in information between speaker and hearer and the purposes which the speaker has at the time of speech. These remarks suggest that Stalnaker does conceive of a comprehensive theory as employing more elaborate representations of context than sets of possible worlds.[8]

II.

The main concern of Stalnaker's paper is the problem of belief attribution. To understand, he argues, how belief attribution works it is necessary to distinguish between two contexts. The first of them, the *basic* context, is the context that identifies the background information presumed by the dis course participants themselves; the second, or *derived* context, identifies what the participants presume about the attitudes of the subject to whom a belief is being attributed. Some complications in the analysis of attitude attributing sentences and utterances arise from the fact that they can rely, in determining the belief they attribute, both on the derived and on the basic context. In particular, a singular noun phrase occurring in the embedded clause of an attributing sentence may contribute to that determination via either context.

Let us, on the basis of the theory I sketched in part I, investigate how these different possibilities may arise in the context of some examples which Stalnaker himself brings up. First a close variant of one he gives:

(8)   Phoebe thinks a prowler got into the vegetable garden last night. She believes he stole her prize-winning zucchini.[9]

Let us assume that the context K_0 in which (8) is asserted contains reference markers for Phoebe and the vegetable garden. A DRS representing such a context would have the form:

(9)

```
|  x              y        |
|  Phoebe (x)              |
|  vegetable garden (y)    |
|      C.(x)               |
|      D (y)               |
```
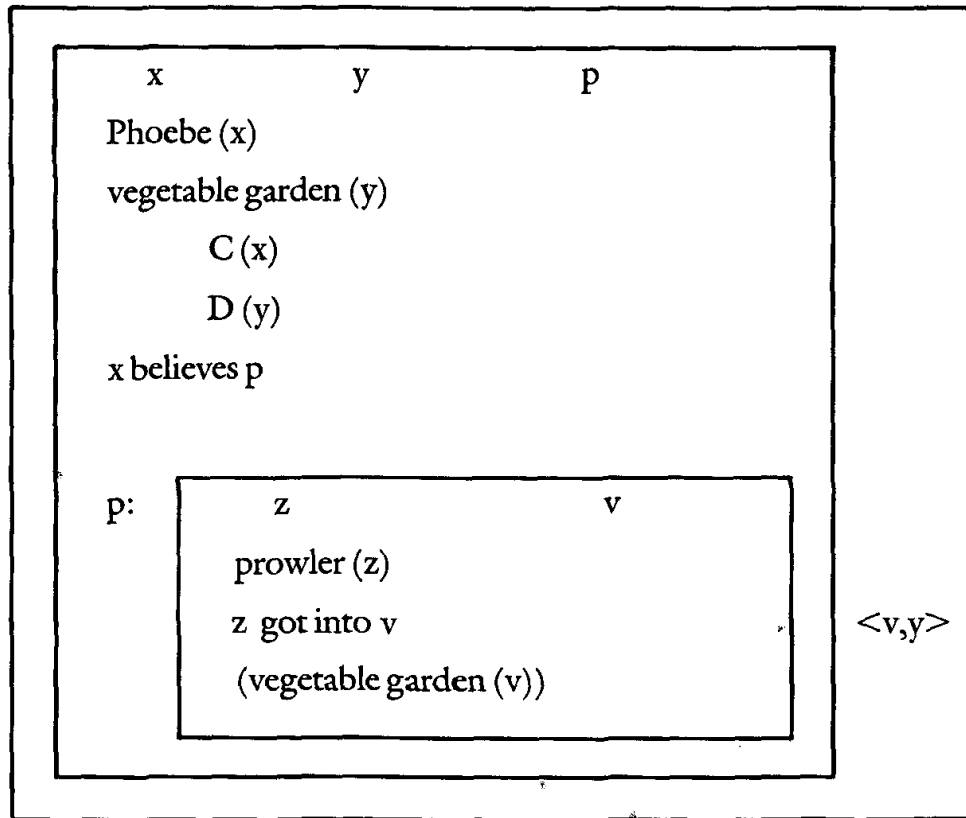
where C(x) and D(y) stand for certain conditions that give more informa-
tion about Phoebe and her vegetable garden but which I will not specify
further. (We may suppose that these conditions uniquely identify the person
and the plot but this is immaterial for what follows.) I will assume that the
conditions C(x) say nothing about Phoebe's beliefs, and thus that, as Stal-
naker would put it, the derived context is initially empty. The condition
'Phoebe(x)' should be understood as requiring that x must be correlated
with the actual bearer of the name, in other words Phoebe. A theory of
proper names might want to say a good deal more about this condition.
But we need not go into that here.

What is the effect of interpreting and accepting the sentence (8.1) in this
context? Taking our clue from what Stalnaker says we should expect the
result to be a pair of contexts, a basic and a derived context. What should
go into the derived context seems, in *some* sense, clear: that should be the
content of the complement clause. But what should the new basic context
be like? Here we encounter a slight dilemma. A recipient who hears and
accepts (8.1) thereby gets a certain bit of information, just as when he hears
and accepts any other meaningful assertion. Of course, what information
he gets depends on the assertion. In the present case it is information that
Phoebe has a certain belief; if the speaker had said instead "Someone got
into Phoebe's vegetable garden last night" it would be the information that
someone got into her garden. But nonetheless, in each of the two cases it
is new information, and thus in some sense it ought to go into the *basic*
context. For it is the basic context which represents the discourse partici-
pants' own presumptions of what is the case. If we want, in our treatment
of what happens when the recipient accepts (8.1), to do justice to this prin-
ciple then the basic context should be modified into one which now con-
tains the additional information that Phoebe believes that a prowler got
into her vegetable garden. The dilemma we encounter is that if this is done
the basic context will already contain the derived context as a part. So we

have here not so much a pair of separate contexts but a single context structure which contains 'another such structure as a subsidiary. While this is somewhat different from what Stalnaker says, it is, I hope, close enough to his intuitions. In any case, this is the way, I think, that the context which (8.1) induces should be represented.

There are still some representational details to be settled. But as the existing options are not important to what I want to say here I will simply present the result as it incorporates the choices I have made, without bothering to give an elaborate justification of all its aspects. Here it is:

(10)

$$
\boxed{
\begin{array}{l}
\boxed{
\begin{array}{l}
\quad\text{x}\qquad\qquad\text{y}\qquad\qquad\text{p} \\[4pt]
\text{Phoebe (x)} \\[4pt]
\text{vegetable garden (y)} \\[4pt]
\qquad\text{C (x)} \\[4pt]
\qquad\text{D (y)} \\[4pt]
\text{x believes p} \\[16pt]
\text{p:}\quad
\boxed{
\begin{array}{l}
\quad\text{z}\qquad\qquad\qquad\text{v} \\[4pt]
\text{prowler (z)} \\[4pt]
\text{z got into v} \\[4pt]
\text{(vegetable garden (v))}
\end{array}
}
\end{array}
}\quad \text{<v,y>}
\end{array}
}
$$

This DRS contains a number of components of types we haven't yet encountered, and a few explanatory comments about their meaning and function are in order before we can proceed further. p is a reference marker for a cognitive state. The condition 'x believes p' represents the information that the cognitive state that p stands for is a state of belief of the individual represented by x. The condition beginning with 'p:' conveys the information that part of the state that p stands for is characterized by the DRS given by the embedded box of (10). This embedded DRS identifies the derived

context. The rules which convert the embedded sentence of (8.1) into the derived context structure are the very same that are used to process non-embedded sentences. (In this sense the account conforms to Davidson's now famous requirement of "semantic innocence," a desideratum which Stalnaker also endorses.)

There is however one complication that arises for embedded sentences but not for non-embedded ones, viz., that constituents of the sentence may have to be interpreted in relation to the basic as well as to the derived context. (10) gives us a first example of this. Definite noun phrases—in particular pronouns, definite descriptions and demonstratives—appear to obey what linguists have called the *familiarity constraint*.[10]

According to this principle the use of a definite NP presupposes that the context already contains an element representing what the noun phrase denotes. (We in fact already relied on this principle once before when we discussed pronouns in Section I.) The principle entails that the construction rules which deal with definite NP's must find a reference marker in the current DRS, and identify the new marker they introduce for the NP to which they are being applied with that first marker. When we apply the rule for definite descriptions to the embedded noun phrase *the vegetable garden* in the context given by (9) there is only one real option, viz., to identify the new marker v in the derived context with the marker y in the "basic part" of (9). This identification is indicated by the ordered pair <v,y> to the right of the embedded box of (10). We will refer to such pairs as *internal anchors*, or simply *anchors* when there is no danger of confusion (we will come to speak later also of *external* anchors). The function of the anchor in (10) is to allow the information which the basic part of the DRS carries about the correlate of y to partake in the determination of the content of the embedded DRS, and thus of the belief which (8.1) attributes to Phoebe. According to Stalnaker this content must be a proposition. (A proposition, to be exact, relative to the derived context; but as by our assumptions the initial derived context was empty, this qualification makes no difference in the present case.) *How* this content is determined, and how in particular the anchor contributes to that determination, is a question I want to leave for the moment. I will return to it below.

There is one further question that the NP *the vegetable garden* raises and that I want to mention briefly. Should the descriptive content of the noun phrase be regarded as part of the attributed belief? In our terminology: Should the embedded DRS be made to contain a condition to the effect that v stands for a vegetable garden (and, perhaps, in addition a condition expressing that, within certain contextually determinable limits, there is only one such)? Or should the descriptive content of the noun phrase be understood as serving no other purpose than to help the interpreter to find

the intended antecedent (the one we have represented by the internal anchor <v,y>)? I have the impression that speakers' intentions vary on this point, and therefore doubt that the question has a clear answer. For all I can see definite descriptions are ambiguous in this particular respect. Sometimes their descriptive content is understood as part of the content of the clause in which they occur, sometimes it is taken to serve solely as a pointer towards the antecedent, and sometimes it is regarded as playing both these roles at once. I have indicated this element of ambiguity in (10) by placing the relevant condition in parentheses.

From what we observed earlier it would seem to follow that the ambiguity we have just noted is only *one* of the ambiguities to which definite descriptions in the complement clauses of attitude verbs may give rise. The other "ambiguity" relates to the question where the antecedent is to be sought—in the basic or in the derived context. In the present case this question has a straight answer, because, as we saw, the derived context does not contain any plausible candidate. In fact, it is worth dwelling briefly on the question in what sense, if any, this disjunction of possibilities—that of locating the antecedent in the derived and that of locating it in the basic context—can be called an "ambiguity." The construction rule for definite descriptions in embedded positions, it is true, should allow for both possibilities. And this may be taken as a justification for saying that the general phenomenon of a definite description occurring in a complement clause is *ambiguous between* them. But this ambiguity, if so it should be called, must be distinguished from yet another, and quite different sense in which a description can be ambiguous. A description can be ambiguous in the particular context in which it is used in that there are at least two candidate antecedents both of which fit its descriptive content. This is a phenomenon that may arise with descriptions in non-embedded clauses just as well as with those which occupy embedded positions, e.g., in

(11)   The accident destroyed both the car's axle and its front fender. The part was irreplaceable.

The part is ambiguous between the axle and the fender. Utterances that are ambiguous in this particular way strike us as defective. Of course, when the description occurs embedded there may be the additional complication that one of the candidate antecedents belongs to the basic and the other to the derived context. But this complication doesn't seem to make the utterance any worse. If it is ambiguous then it will be deviant irrespective of whether the candidate antecedents belong to the same part of the context or to different parts. Or, to turn the point around, if the utterance *is* felicitous then any definite descriptions it contains will point unequivocally towards a unique contextual element. In which part of the context that element can be found is a further, and different, question.[11]

When a definite description in embedded position is used felicitously then the question whether it contributes via the basic context or directly via the derived context itself must have a definite answer. For indefinite descriptions this is in general not so. Consider for instance the subject *a prowler* of the embedded clause of (8.1). The particular treatment of this phrase that led to the DRS (10) assigned it what in traditional terminology would be described as a "narrow scope" reading: The belief that is attributed to Phoebe is the notional belief that "some prowler or other" was in her garden. However, for indefinite descriptions in complement clauses of attitude verbs there is in general also another, "wide scope" interpretation, which introduces the marker into the basic context. Had we processed the noun phrase in this other way we would have obtained a DRS to the effect that Phoebe has, concerning some particular prowler, the belief that he got into the vegetable garden.[12]

Let us now consider the second stage of the interpretation process, which concerns (8.2). The first step deals with the pronoun *she*; the result, as described in Section I, is that of identifying the new marker, w, say, which the pronoun introduces with the appropriate marker in the basic part of (10). For reasons of gender this can only be the marker x. The next step must deal with the verb phrase (or more accurately with an intermediate condition of the form 'w believes that he . . .'). In fact, the treatment of (8.1) involves a similar step. That step I chose to ignore, but here it is good to pause briefly and to note that to perform the necessary operations the algorithm will have to register the synonymy or near-synonymy of *think* and *believe* and the identification of w and x. For only on this basis will it be able to conclude that it is the derived context of (10) which the embedded clause of (8.2) should modify, and not the basic context (or, a third possibility, some new derived context to be set up in addition to the one (10) already contains). We won't worry here about the rule or rules responsible for this step (any more than we have about the details of all the other rules of which I have already made implicit use). But it is worth registering that when we proceed to the study of attributions that involve other attitude verbs besides *believe*, or involve more than one attitude subject, these rules become surprisingly delicate and complex.

The remaining steps of the procedure interpret the embedded clause of (8.2) relative to the derived context of (10). The first of these deals with the pronoun *he*. This definite NP finds, in contrast to the definite description of (8.1), its antecedent in the derived context. Indeed, this step can be executed in precisely the way in which we dealt with anaphoric pronouns in Section I.
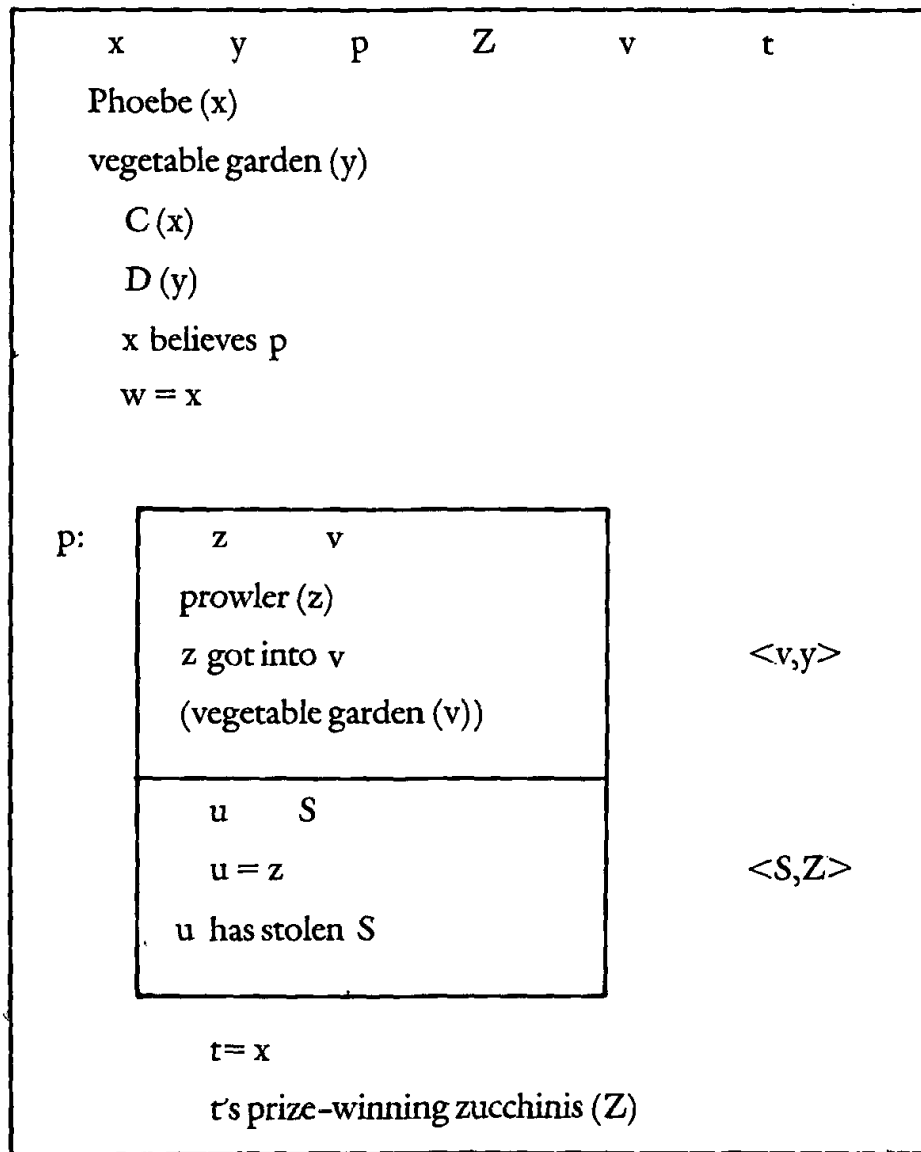
The last step must deal with *her prize-winning zucchini*. Here we face yet another problem, for neither basic nor derived context contains a marker that could serve as antecedent for this definite noun phrase. This is a situation which violates the familiarity constraint as I have stated it. In such

cases, or so some familiarity theorists would have it, the interpretation procedure must rely on a form of *accommodation*, in the sense of Lewis.[13] The accommodation needed to compensate for a violation of the familiarity constraint would take the form of modifying the context into one which incorporates the information that there is a unique individual satisfying the content of the description. In the present case this could be done in two different ways: the accommodation can take place either at the level of the basic or at that of the derived context. The first option means that the recipient commits *himself* to the existence of a (unique) set of prize-winning zucchinis that belong to Phoebe, the second that he imputes this assumption only to Phoebe. I have chosen the first option.

Putting it all together we obtain for the full passage (8) the DRS

(12)

x	y	p	Z	v	t

Phoebe (x)

vegetable garden (y)

   C (x)

   D (y)

   x believes p

   w = x

p:

z	v

prowler (z)

z got into v   <v,y>

(vegetable garden (v))

u	S

u = z   <S,Z>

u has stolen S

t = x

t's prize-winning zucchinis (Z)

This DRS has a number of components we did not encounter in Section I and it is not evident from what I have so far said how these affect the truth conditions. But I want to hold off a little on this point, and first have a look at the other example which Stalnaker discusses in some detail. This is the well-known case of Ortcutt and Ralph, of which Stalnaker notes that it presents the same problems as that of the hapless Pierre[14] on which so many of the contemporary discussions concentrate. When we look at the phenomenon of belief attribution in the way Stalnaker proposes then the Ortcutt example—and the same is true for all related examples that can be found in the recent literature—is a good deal more complex than many earlier discussions of the problem made us realize. From the new perspective it is clear that we cannot simply ask whether the attributions

(13.a) Ralph believes that Ortcutt is a spy.

and

(13.b) Ralph believes that Ortcutt is not a spy.

are true.

Indeed, that is a question that Stalnaker does not even explicitly address. The first question that we have to ask is: How do these sentences modify the context in which they are used?

This is a question we cannot answer until we are told what that context *is*. Obviously there are various possibilities here. Stalnaker notes this, though only in passing. But it·will be useful, I believe, to go into a little more detail than he provides.
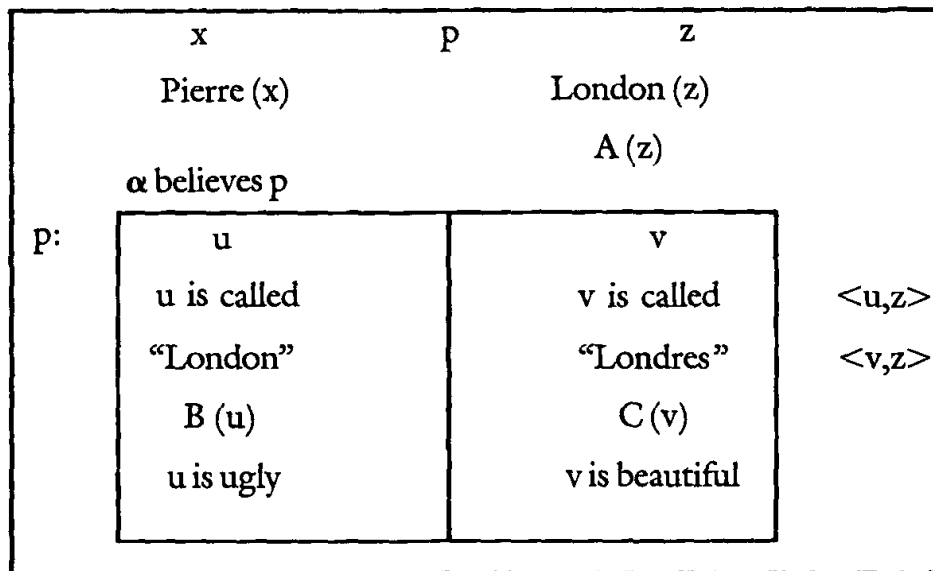
The case about which Stalnaker is most explicit is that of a context whose derived part represents Ralph as having two separate conceptions of Ortcutt. How should such a context be rendered? Here we face once again the problem of connection between basic and derived context. As Stalnaker describes the case there are in Ralph's presumed belief worlds two distinct individuals each of whom has a claim to being the real Ortcutt, but only one of them could be the real Ortcutt, and it is undecided which one that is. Suppose now that within this context someone makes one of the attributions listed under (13), say (13.a). Suppose also that the only way in which the name "Ortcutt" in the complement clause of (13.a) may be interpreted is that according to which it contributes to the content of the attributed belief via the basic context. (More on this point below.) Then, so Stalnaker, the interpretation process must abort, because it cannot be decided which of the two individuals of the derived context is the real Ortcutt: There is no unambiguous route from the intermediate station in the basic context to a final destination in the derived context.

I have problems with this solution, for one because I have difficulty understanding what we should make of the assumption that it is open which of the two individuals in Ralph's belief worlds is the real Ortcutt, and for

another because it seems to me that even if there are cases which might be described in the way Stalnaker does describe his case there are others where this description does not seem at all plausible and where for that reason his analysis of the problem would not apply anyway. To lead up to what I have in mind let me recall some of the things Kripke says about Pierre. Pierre, he tells us, has formed in his childhood a belief he then expressed, and may still express, as "Londres est jolie." And given the way in which he acquired the belief it seems entirely in order to construe it as a relational belief which consists in his attributing beauty to the city of which his friends have spoken to him. Indeed, had Pierre never *left* France then probably no problem about that way of construing his belief would ever have arisen. Later he forms a second belief—one he expresses, in his second language, as "London is ugly"—about the place in which he then actually lives and thus knows from immediate experience. That this belief is also to be construed as a relational belief appears even less controversial.
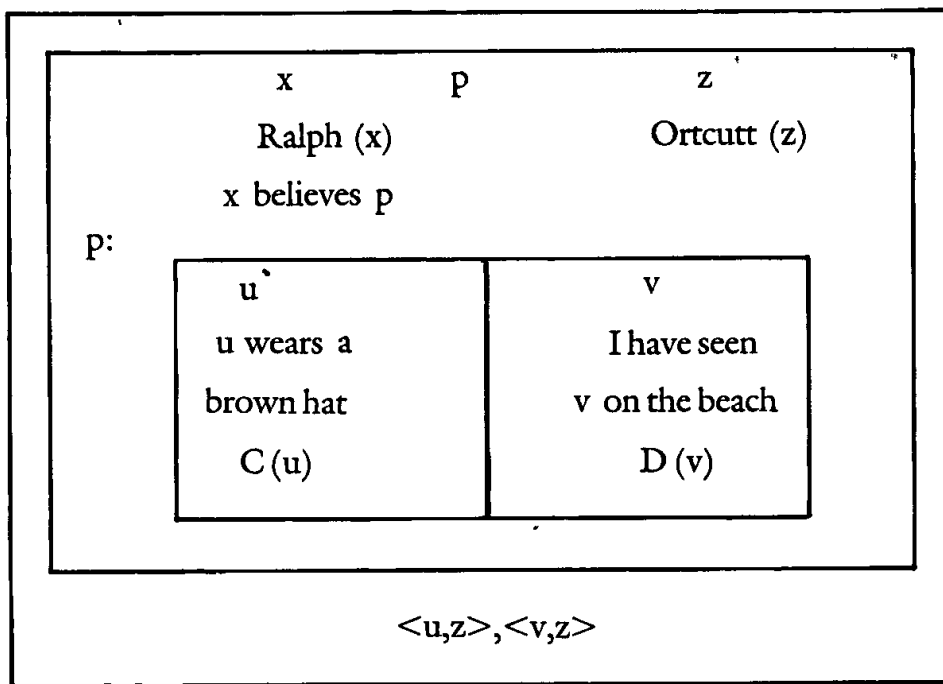
Controversial or not, there certainly are people who on the strength of this story would construe Pierre's beliefs as relational beliefs. For such people the context in which they are to interpret further belief attributions to Pierre will connect both the individual in Pierre's belief worlds which he knows as "London" and the one he knows as "Londres" to the London of the basic context. A DRS representing the way such a person would understand Pierre's beliefs will have an anchor for each of the two markers that in the derived context stand for the place Pierre knows under the name "London" and for the one he knows under the name "Londres," and both these anchors will anchor the marker from the derived context to the *same* marker of the basic context, which stands for London. In other words, it will look something like this:

(14)

x	p	z	
Pierre (x)		London (z)	
		A (z)	
α believes p			

p:

u	v	
u is called	v is called	$\langle u,z\rangle$
"London"	"Londres"	$\langle v,z\rangle$
B (u)	C (v)	
u is ugly	v is beautiful	

This structure reveals, by the way, one point that looms large in Kripke's discussion of the case, viz., that in some sense Pierre's beliefs are incompatible. The anchors indicate that from the perspective of the basic context the two parts of Pierre's belief state which are explicitly represented are incompatible; for they attribute incompatible properties to what is, from this perspective, one and the same individual, viz., the one that z stands for. But from Pierre's own perspective, which is reflected in the embedded DRS, there is no such incompatibility, for the information that u and v stand for the same city is not part of it. Let us return to Ortcutt and Ralph. Someone who is being told that story is no less justified than he would be in the case of Pierre to construe Ralph's beliefs as relational beliefs. Accordingly the DRS representing the contextual background against which such a person will interpret additional belief attributions to Ralph will have a form analogous to that of (14):

(15)

$$
\begin{array}{|l|}
\hline
\quad \boxed{\begin{array}{ll}
\qquad x \qquad\qquad p \qquad\qquad z \\[4pt]
\qquad \text{Ralph (x)} \qquad\qquad \text{Ortcutt (z)} \\[4pt]
\qquad x \text{ believes } p \\[6pt]
p: \quad \boxed{\begin{array}{l|l}
u & v \\[4pt]
u \text{ wears a} & \text{I have seen} \\[4pt]
\text{brown hat} & v \text{ on the beach} \\[4pt]
C\,(u) & D\,(v)
\end{array}}
\end{array}} \\[8pt]
\qquad\qquad \text{<u,z>,<v,z>} \\
\hline
\end{array}
$$

Someone who has to interpret (13.a) in this context finds himself in a predicament not unlike the one Stalnaker describes. Again, if the interpreter takes the name as contributing via the basic context he will be at a loss what this contribution is. This time the dilemma arises because the context contains two conflicting pointers, rather than a single one that is equivocal. But the effect is the same.

So far I have gone along with Stalnaker's assumption that the name "Ortcutt" must contribute via the basic context. But in fact I do not see that

this is always necessary. Perhaps the version of the story Stalnaker has in mind is one where·this *is* the only option. If for instance Ralph knows neither of the man on the beach nor of the man in the brown hat that he is called "Ortcutt," and if this additional information is part of my contextual background, then I am in a position to conclude that the name "Ortcutt" can contribute to the belief which (13.a) attributes only via the basic context. But the case need not have been this way. And the case of Pierre is definitely otherwise. Pierre knows the two cities in his belief worlds by name. (It is just that he doesn't know them by all their names.) Thus, in the context characterized by (14) a new attribution, e.g.,

(16) Pierre believes that Londres has many beautiful parks.

permits, it seems to me, the following interpretation: The name "Londres" is taken to contribute via the derived, not the basic context. In the derived context, i.e., the embedded DRS of (14), there is exactly one marker that fits the name—in the sense that according to the context it stands for something which bears the name—and it is to this marker that the predicate of the embedded clause must therefore be attached. In fact, precisely when the alternative strategy, which attempts to interpret the name via the basic context, fails because of ambiguity in the connection between basic and derived context, the strategy just sketched will present itself as the only sound one. It is for such reasons that the attributions

(16.a) Pierre believes that London is beautiful.

and

(16.b) Pierre believes that Londres is ugly.

seem, against the background of Kripke's story, so much worse than

(16.c) Pierre believes that London is ugly.

and

(16.d) Pierre believes that Londres is beautiful.

These few observations indicate how complicated the matter of belief attribution really is, even for comparatively simple sentences and contexts. Note in this connection that, to return once more to Ralph, we have considered only one of the many different contexts in which (13.a) and (13.b) could be used. In a context, for instance, whose derived part has a marker for the man in the brown hat but not for the man on the beach, (13.a) will produce an acceptable interpretation. If, on the other hand, the derived context only has the man on the beach, the very same statement will be highly misleading.

It would be possible to go through more variants of the Ortcutt case. But what has been said should be enough to illustrate in what manner the significance of the attributing sentence may be affected by the context in which it must be interpreted. This is one of the reasons why the "change of perspective" which Stalnaker advocates in his paper is both necessary and illuminating.

However, a change in perspective cannot make all the old questions simply disappear. There is one question, in particular, which has often been taken to be the central question that a semantics of belief attribution should address, but upon which we have only barely touched. This is the question: When is a belief attribution true? No theory of belief can afford to ignore this question altogether. And so I feel I cannot leave the subject without at least raising the matter.

For Stalnaker the answer would, I expect, come to something like this. Believers are in certain belief states and these states determine sets of possible worlds, their *belief sets*. Suppose that a speaker utters in a given context C the sentence "A believes that α," and that the derived context of C is correct in the sense that it determines a set of possible worlds in which the actual belief set of A is included. Then the attribution is true if the modified derived context still determines a set of worlds that includes the belief set. (This leaves open what we are to say of attributions that are made against the background of contexts that misrepresent A's beliefs. But this is not necessarily a defect. For it is in fact often quite unclear what to make of an attribution that is made in such a context.)

For Stalnaker, who takes contexts to be given as sets of possible worlds, this definition is, as far as I can see, unproblematic. But for the treatment I have proposed no comparable definition is as yet available. To provide one I must begin by doing what I promised earlier, viz., to state the truth conditions for the DRS's we have been constructing in this section. As a matter of fact, to articulate the truth conditions of such DRS's would be simple enough if we could treat them essentially as pairs of basic and derived contexts. Then the truth definition mentioned in Section I could be extended by adopting the following stipulations (adumbrated by what I said about the corresponding DRS conditions above):

(i)   a condition 'x believes p' is satisfied if the correlate of p is a belief state of the correlate of x;

and

(ii)   a condition 'p:K' is satisfied if the set of worlds determined by the DRS K includes the set of worlds determined by p.[15]

This is simple enough. But it is incomplete. One aspect has been left out, for we haven't said anything about the anchors. So far I have only talked

about how anchors get introduced and about the role they play in the interpretation of noun phrases in embedded clauses. But surely this can't be all there is to them./The presence of an anchor means, intuitively that the parts of the derived context in which the anchored element occurs represent relational beliefs. If the anchor were absent they would represent beliefs that are not relational, but notional. Thus the presence of an anchor·in a DRS makes a difference to its meaning, and we should expect the truth definition to reflect this.

So *what* is the difference that anchors make to truth conditions? From what has been said about them already the, answer might seem straightforward. Intuitively an anchor conveys that the markers it connects are to be considered as standing for the same thing. In truth-conditional terms this simply means that the only admissible correlation of individuals with markers must correlate any anchored marker to the same individual as the marker with which the anchor connects it. Simple as this may be from a formal standpoint, the intuitive significance of this addition to the truth definition is not so straightforward. To see its implications let us assume that there is such a thing as relational belief:[16] that a believer can have about an individual b the belief that it is a P, i.e., have a belief whose content is the singular proposition that b has P. A DRS characterizing such a belief must make explicit that the relevant reference marker can *only* be correlated with the individual that the belief is about. The device which has been used in DR theory for this purpose is that of an *external anchor*. An external anchor is a function that maps reference markers onto real-objects (in contrast, internal anchors, as we saw, link reference markers with other markers). The effect of the anchor is to constrain the possible correlations between markers and objects: every anchored marker may be correlated only with the object onto which the anchor maps it. To give an example, if Phoebe believes about the biggest zucchini in her garden, call it b, that a prowler stole it, the content of that relational belief is captured by the *anchored* DRS

(17)

$$\left[\begin{array}{cc} z & u \\ \text{prowler } (z) \\ z \text{ has stolen } u \end{array}\right] , \ \{<u,b>\}$$
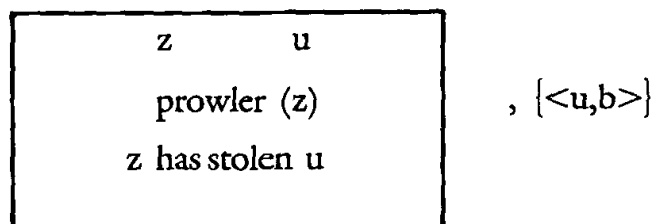
If we accept that there is such a thing as relational belief we must also recognize that different beliefs may be about the same things and in particular that different people can be related in belief to the same object. For

instance, several people may have beliefs, though not necessarily the same beliefs, about Ortcutt. Suppose that speaker and recipient of (13.a) are among them, and that the contextual background against which (13.a) is communicated contains the relational ·beliefs they share about Ortcutt. Then, in the first place, the marker which stands for Ortcutt in the basic part of the context should be externally anchored to Ortcutt. Suppose furthermore that speaker and·hearer understand Ralph to be also related in belief to Ortcutt. Then the corresponding markers in the derived context should be externally anchored to Ortcutt as well. To account for·these anchors explicitly we need an anchored DRS, say the one we obtain by adding them to the DRS (15). In this anchored DRS we can see the internal anchors, which connect the markers u and v in the derived part with the marker z in the basic part, as an internal "reflection" of the fact that all three markers are externally anchored to the same real individual: The constraint which the internal anchors impose will be automatically satisfied by any correlation which obeys the external anchors.

If I felt we could assume that all internal anchors are reflections, in the sense explained, of external anchors, I would be content to leave the question of truth at this point. But I do not think we can assume this. I can construe your beliefs as being about a thing that I have certain beliefs about myself, even if in reality there·is no thing that either your beliefs or my beliefs are about. And a third person who knows there is no such thing may still recognize our beliefs as being, in some sense, about the same object. (It is this construal of others' beliefs that motivates and justifies such attributions as Geach's famous "Hob-Nob-sentence.") Contexts representing such belief construals will contain internal anchors without corresponding external anchors.

Here the constraint the internal anchor imposes—that the markers it connects must always be correlated with the same individual—can no longer be explained in terms of causal connections between believing minds and external objects. Rather the explanation must be·in terms of a relation between the different subjects to whom the beliefs are being attributed. The justification of the internal anchor must be something like a *joint commitment* on the part of those subjects, or, in other cases, a collective commitment made by the community to which they belong—the commitment to *treat* the respective beliefs *as* being about the same thing. We arrive at such joint commitments in part by talking to each other, and by taking ourselves on those occasions to be talking about the same things. And the agreements we inevitably enter into on these occasions are meaningful because we live in one and the same world, and face the same spectrum of external objects, among which the instantiations of our beliefs must be found if they are correct.

Once we discern this kind of intersubjective commitment as an inaliena-

/

ble ingredient of human communication and cognition we are in a position to interpret internal anchors as the reflectants of such commitments. This is indeed the role I intended for them. (So acceptance of the version of DRT I have outlined in this section entails that one accepts this concept of inter-subjective commitment as genuine.) This implies however that the meaning of an internal anchor will never be one that an external anchor could carry in its stead. Thus éven in a situation like the one discussed earlier, where the external anchors render the internal anchors superfluous to the truth condition of the structure, the latter still have a separate contribution to make.

The irreducibility of internal to external anchors is manifest in the princi-ples that govern attribution interpretation. There are situations we saw, in which a given definite NP can make a particular contribution to the derived part of the context only when an internal anchor is present which mediates between the basic part of the context and its derived part. External anchors cannot provide this kind of mediation. To see why there should be this dif-ference between internal and external anchors it is necessary to remember that the contextual information which affects the way an utterance is inter-preted must be available to the interpreter. For after all, it is he who is doing the interpreting. Now, external anchors do not represent information that is available in the sense relevant here. We already got a glimpse of this in connection with Pierre and Ralph. Each of them is in a belief state represent-able by a DRS in which two markers, u and v, say, are externally anchored to the same individual. But the information that the markers are anchored to the same individual is something neither Ralph nor Pierre has. There are other instances of this predicament in which one of the beliefs—say, one involving the marker v—is to the effect that some other person, call him b, has a relational belief concerning the individual represented by v. In this case the information that u and v stand for the same individual would, in the DRS's we have been using, be represented by the internal anchor that links v to u. Someone whose information state is of this latter kind may be at a disadvantage when he confronts new belief attributions to b. He may be unable to assign to the attribution the interpretation that he would have been able to give it but for the information that u and v stand for the same thing.

An example may help to clarify what I have in mind. Suppose that you are staring at a man. This man is in fact Ortcutt but you don't know this. All you think is "I would not care to meet that man in the dark." Surely, if there exists such a thing as relational belief at all, your thought is a relational belief about Ortcutt. You may also know that your friend Ralph has a high opinion of some person whom he has met personally and of whom he has told you various things, referring to him in the process as "Ortcutt." So, in

fact, you and he have beliefs about the same person. In the DRS represent-ing the situation as you see it there will be a marker which stands for the man at whom you are staring and a marker that stands for the man called "Ortcutt" inside the embedded DRS which represents the beliefs that you take Ralph to have; and both these markers will be externally anchored to the same Ortcutt.

Suppose that I, who am standing next to you, who recognize the man as Ortcutt, and who assume that you realize him to be Ortcutt too, say to you,

(18)   Ralph thinks the world of that man.

In reaction to this you may come to doubt Ralph's judgment of men. But you won't be able to interpret my statement as attributing to Ralph beliefs about someone that you already know him to have other beliefs about. If instead I had said,

(19)   That is Ortcutt. Ralph thinks the world of him.

you would have been able to make the connection that you can't make on the strength of the first report. You could have made the connection in the second case because the first sentence of (19) would enable you to set up the internal anchor that provides the needed link between the man before you and the man you know Ralph to know. But in response to (18), where nothing helps to establish this internal anchor, the external anchors will be powerless to produce in you that understanding of my statement which I intended.

This last discussion, of the semantic import of internal anchors and what might be called their "socio-psychological foundation," is one which goes beyond anything that Stalnaker discusses. As such it may seem misplaced in a commentary on his paper. My justification for including it is that unless we take this aspect of meaning and use into account it does not seem possi-ble to explain what belief attributions mean by analyzing how they are inter-preted in context. It ought to be evident from what I have said in these comments how much I am persuaded by the general approach which Stal-naker has taken. If I have added something that goes beyond the substance of his argument it has been only in the hope that that may contribute some-thing to the further exploration of the new ground he has broken.