SITUATIONS AND ATTITUDES*

READING the early work in logic by Frege and Russell, one can hardly fail to be struck by the extent to which their thinking was shaped by a concern to understand the verbs of cognitive attitudes, verbs like *wonder, believe,* and *know*. In spite of this concern, and all the subsequent progress in logic, there is still no satisfactory systematic account of the logic of the attitudes.

In this paper we outline an approach which we believe will lead to a satisfactory and systematic account. For our sample of verbs we take *see, know, believe,* and *say*. We call our theory *situation semantics;* it is closer in spirit to Russell than it is to Frege in some fundamental ways. We begin with some key features of situation semantics, and then move to a discussion of philosophical and semantical issues surrounding the attitudes. A rigorous semantics for a fragment of English incorporating these verbs as well as tense, indexicals, demonstratives and definite descriptions, proper names, pronouns, and conditionals is in preparation.

SITUATIONS

Situations are basic and ubiquitous. We are always in some situation or other. Human cognitive activity categorizes these situations in terms of objects having attributes and standing in relations to one another at locations—connected regions of space-time. Human languages reflect (and enhance) this cognitive activity by giving us a way of communicating information about situations, both those we find ourselves in and those removed from us in space and time.

In attempting to develop a theory of linguistic meaning that concentrates on situations, we recognize the epistemological primacy of situations, but follow the lead of language and take ob-

*To be presented in an APA symposium on The Logic of Perception and Belief, December 30, 1981. Richmond Thomason will comment; his paper is not available at this time.

This paper represents joint work of the authors; the order of names is merely alphabetic. The first author is grateful to the National Science Foundation. The paper was completed while the second author was a fellow at The Center for Advanced Study in the Behavioral Science, and he is grateful for support from the Center, the National Endowment for the Humanities, the Andrew Mellon Foundation, and Stanford University. Both authors are grateful to the Center for gracefully accommodating the needs of collaborators.

We owe much to discussions with John Etchemendy, Michael Turvey, and others at the Center and at Stanford. Our point of view was profoundly influenced by Turvey and others working in the tradition of ecological realism. For an introduction to this point of view, the reader may consult Claire Michaels and Claudia Carello, *Direct Perception* (Englewood Cliffs, N.J.: Prentice-Hall, 1980).

0022-362X/81/7811/0668$02.40

jects, relations, and locations as the primitives of our theory, reconstructing situations from them. Thus we have as primitives:

(i) a set A of individuals $a, b, c \ldots$;
(ii) a set R of relations, $R = R_0 \cup R_1 \cup \ldots R_n \cup \ldots$, where R_n consists of the n-ary relations; and
(iii) a set L of space-time locations $l, l_1, \ldots$

A *situation* s is characterized by its location l and its *type* s, $\mathbf{s} = \langle l, s \rangle$. The type represents which objects stand in which relations at the location. We represent these types by means of partial functions from relations $r \in R_n$ and sequences $\langle a_1, \ldots, a_n \rangle$ of objects to 1 (true) and 0 (false).[1] The partial function s_0 defined by

$$s_0 \text{ (awake, Jackie)} = 1$$
$$s_0 \text{ (awake, Molly)} = 0$$

will be realized in those situations **s** where the first author's dog is awake, the second's asleep, regardless of what the reader's dog is doing, if she or he has one. (s_0 is realized in $\mathbf{s} = \langle l, s \rangle$ if $s_0 \subseteq s$.) We use S for the set of situation types $s, s_0, s_1, \ldots$ and **S** ($= L \times S$) for the set of situations **s**, $\mathbf{s}_0, \mathbf{s}_1, \ldots$

A *course of events* σ is a partial function from the set L of locations into S. Thus every course of events is also a set of situations, at most one at any given location l. If $l \in \text{domain}(\sigma)$ we write σ_l for the situation type $\sigma(l)$. We use Σ for the set of all courses of events. A *total course of events* is a course of events defined for all locations. We distinguish one among these as the *actual course of events* σ^*. A situation $\mathbf{s} = \langle l, s \rangle$ is *actual* if $s \subseteq \sigma_l$; that is, if the type of **s** is part or all of what is actually the case at l.

A (realistic) *proposition* is a set $P \subseteq \Sigma$ satisfying:

(Monotonicity) $\sigma \in P$ and $\sigma \subseteq \sigma'$ implies $\sigma' \in P$.

The adjective 'realistic' here is used to emphasize that these are constructs of real objects, properties, and locations, not things in someone's head. [A (realistic) unlocated proposition is a set $P \subseteq S$ satisfying a similar monotonicity constraint: $s \in P$ and $s \subseteq s'$ implies $s' \in P$.]

There are three basic relations on space-time locations that are represented in English:

$l_1 \circ l_2$	l_1 temporally overlaps l_2
$l_1 < l_2$	l_1 temporally and wholly precedes l_2
$l_1 @ l_2$	l_1 spatially overlaps l_2

[1] We regard truth values as slipping into the universe in the process of abstraction from situations to objects standing or not standing in various relations.

We take these to be extensional relations on L, relations out of which one can construct "instants" of time and "points" of space in the manner of Whitehead and Russell.

The starting point of situation semantics is that untensed indicative statements describe or designate situation types and that tensed indicative statements designate propositions, sets of courses of events. We use the term *statement* advisedly, here, for a sentence like *I am sitting* can be used to make as many different statements as there are speakers and times to utter it. The sentence has a fixed "meaning," but the different statements will describe different events. That is, the different statements will have different "interpretations." This distinction between meaning and interpretation is the subject of the next section.

MEANING AND INTERPRETATION

A number of important themes in situation semantics can be developed by discussing the following simple sentences:

(1) *I am sitting.*
(2) *Sandy is sitting.*
(3) *She was sitting.*

Let us begin with the word *I*. A reasonable thing to say about this expression is that, whenever it is used by a speaker of English, it stands for, or designates, that person. We think that this is all there is to know about the meaning of *I* in English and that it serves as a paradigm rule for meaning.

Consider the relation:

The expression α (of the language L) as used by x, stands for y.

which we write as $[\![\alpha]\!]\,(x, y)$. A theory that tells us every condition under which $[\![\alpha]\!]\,(x, y)$ holds is our candidate for a theory of meaning for the language L. What we were told about *I* gives us one condition:

$$[\![I]\!]\,(a, y) \text{ iff } a = y$$

This relation view of meaning demands that systematic attention be paid to the appropriate values of each coordinate. Our starting point in situation semantics is that, when the first coordinate α is a tensed indicative sentence, then courses of events are the appropriate third coordinate y.

This decision has ramifications for the second coordinate x. It shows that speakers are too simple a choice for this coordinate. Sentence (1) can be used by a single person at different places in space-time to describe different events. Similarly the designation of

you, now, she, this, was varies from utterance to utterance, depending on who the speaker is talking to when, and about whom, what, and when. We represent the utterance-specific facts with reference to *discourse situations* and *connections.*

A *discourse situation d* represents the situation in which the speaker and addressee find themselves. It consists of a situation $s_d = \langle l_d, s_d \rangle$ with a designated individual a_d such that $s_d(\text{speaks}, a_d) = 1$. We modify the rule for *I* given above to

$[\![I]\!]\ (d, y)$ iff $y = a_d$.

Similarly, *now* constrains the time being referred to overlap the time of utterance; so we can define

$[\![now]\!]\ (d, y)$ iff $y \in L$ and $y\ \mathbf{O}\ l_d$.

Similarly,

$[\![here]\!]\ (d, y)$ iff $y \in L$ and $y\ @\ l_d$.

However, there are often utterance-specific facts that have constituents not present to the actual discourse situation. Consider sentences (2) and (3) above, for example. It is reasonable to suppose that in an interpretable utterance of (2) [or (3)] *Sandy* stands for Sandy (or that *she* stands for some female). But which Sandy—Sandy Koufax, Sandy Dennis, or Little Orphan Annie's dog? What is unaccounted for here is that a meaningful use of (2) is about some specific individual Sandy [and that a meaningful use of (3) is about some specific female]. Since these individuals need not be present in the actual discourse situation, we have no choice but to recognize another component of our second coordinate, a component representing the connections *c* between certain words and things in the world implicit in any meaningful use of those words. Thus an utterance of (2) where the speaker was talking about Sandy Koufax would be represented by the expression (2), a particular discourse *d*, and a partial function *c* with $c(Sandy)$ = Sandy Koufax. We can then represent the meanings of *Sandy* and *she* by

$[\![Sandy]\!]\ (d, c, y)$ iff $c(Sandy) = y$ and y is named *Sandy*.
$[\![She]\!]\ (d, c, y)$ iff $c(she) = y$ and y is a female.

(Even this is overly simple, but it is good enough for now.)

We have now disposed of the noun phrases in (1) to (3), and have the tools at hand for disposing of the verb phrases *am sitting, is sitting,* and *was sitting.* These are all various progressive forms of the verb *sit.* Like most verbs, its interpretation is d.s.i.—insensitive to the discourse situation in which it is uttered. However, *sit* can be

used to designate either an activity $\text{sit}_a \in R_1$, the activity of sitting down, or a state $\text{sit}_s \in R_1$, the state of being seated. It's up to the speaker. Thus connections come up again:

$[\![sit]\!](d, c, y)$ iff $c(sit) = y$ and $y = \text{sit}_a$ or $y = \text{sit}_s$.

Now we turn to the tense of (1) to (3). Like *now*, the present-tense forms of (1) and (2) indicate that the sitting is taking place at a time that temporally overlaps the time of the utterance. The past-tense forms are used to indicate that the sitting took place in the past. But, just as part of the meaning of *she was sitting* is that it is must be used of a particular female to make a statement, so too to make a statement it must be used of a particular past space-time location. To interpret correctly my claim that she was sitting, you must correctly interpret my uses of *she* and *was* as being about a female and a past space-time location. To represent the connections between tense markers and space-time locations, we allow our connections to assign space-time locations to tense markers. Thus:

If α is *am/are/is*, then $[\![\alpha]\!](d, c, y)$ iff $c(\alpha) = y \in L$ and $y \circ l_d$.
If α is *was/were/was*, then $[\![\alpha]\!](d, c, y)$ iff $c(\alpha) = y \in L$ and $y < l_s$.

When we fix all that is specific to a particular utterance of an expression α we obtain what we call the *interpretations* of the utterance. Thus if we fix a particular expression α, discourse situation d, and connection c, we obtain those y such that $[\![\alpha]\!](d, c, y)$ holds, which we write alternatively as $y \in {}_{d,c}[\![\alpha]\!]$. If there is a unique such y, we call y *the* interpretation of the utterance α,d,c and write ${}_{d,c}[\![\alpha]\!] = y$. Thus, e.g., ${}_{d,c}[\![I]\!] = a_d$ and ${}_{d,c}[\![was]\!] = c(was)$, a certain location $l < l_d$.

We can now assign meanings to all sentences α of the form

NP PROG VP

where NP $\in$ {*I, Sandy, she*}, PROG $\in$ {*am, are, is, was, were*}, and VP = *sit*, namely:

$[\![\alpha]\!](d, c, \sigma)$ iff $\sigma_l(P, a) = 1$

where ${}_{d,c}[\![\text{PROG}]\!] = l$, ${}_{d,c}[\![\text{VP}]\!] = P$, and $[\![\text{NP}]\!] = a$. An utterance of α describes a course of events σ just in case a is sitting (in the appropriate sense) at the intended location l in the course of events σ. Notice that ${}_{d,c}[\![\alpha]\!]$ is a proposition, a monotone set of courses of events.

The importance of the meaning/interpretation distinction for an understanding of the attitudes cannot be overemphasized. It rests in part on the following two related facts—(1) efficiency and (2) perspective-relativity.

Efficiency: A given expression α with a single meaning $[\![\alpha]\!]$ can be used in different circumstances with different interpretations.

A word like *I*, for instance, can be used to designate any of us. Although this makes language efficient, allowing a given expression to be used over and over to different ends, it also has a corollary. A sentence that describes a given situation from one person's perspective won't in general describe the same situation from some other perspective. In order that we may get at the same situations, a human language will satisfy the following principle:

Relativity: Different expressions with different meanings can be used in different circumstances with a single interpretation.

Thus, for you, the future reader, to express the fact that I am sitting (now), you could say "He was sitting."[2]

One might say that meaning is a function from discourse situation and connections to interpretation. Thus the proposition, $_{d,c}[\![\alpha]\!]$ is the unique set of courses of events σ such that $[\![\alpha]\!](d, c, \sigma)$ holds. This isn't wrong, but it can be very misleading. The interpretation, the set we get when we fix expression, discourse situation, and connections is very important. It amounts to one sort of uniformity over utterances, and it is a very important uniformity; recognition of its importance is built right into language. But by remembering that meaning is a relation we are reminded of a number of other important uniformities, and these are crucial in understanding the attitudes.

Besides the interpretation, we can construct a number of "inverse interpretations," fixing the last coordinate of meaning and allowing the others to vary. Such inverse interpretations are used in daily life, and are often implicitly involved in talk of "truth conditions" and "when a sentence is true." Suppose, for example, that we say that a child *b* understands *this is milk*, because she says it only when it is true. What we mean is that she says it only when attending to a glass of milk. We are appealing to

$$\{\langle d, c\rangle\} \mid [\![\textit{this is milk}]\!](d, c, \sigma^*) \text{ and } a_d = b.$$

And when we think that this is a good test for understanding this sort of sentence, it is because we think there is some uniformity, perceptually discoverable by a_d, across this set or a significant

[2]The astute reader will realize that in describing my situation, his or her connections are:

$$c(he) = \text{me} \qquad c(was) = \text{here and now}$$

These are objectively determined by which of the authors wrote this part of the paper, where and when, and are independent of the reader's ability to specify them in some more complete fashion. On the other hand, the reader's reading is connected, causally, with the writing.

subset of it. The uniformity is not interpretation, assuming that she interacts with different bottles of milk at different times.[3]

Some philosophers hold that the true vehicles of meaning can be neither relative nor efficient; this view leads to logical atomism. Others think the true vehicles can be relative, but not efficient. This leads to the view that the efficient sentences of natural language must be backed by senses or mental representations that take up the slack—that are "complete in every respect." We think that sentences of natural language are true vehicles of meaning, and that the slack is taken up by other factors in the utterance. Language learning requires coordination of language with the more and less remote parts of the environment, not with senses or mental representations.

INNOCENT ATTITUDES

Statements made with sentences of the following sort we call *attitude reports:*

(4) *Agnes saw me jump in the fountain.*
(5) *Agnes saw that I was sitting in the fountain.*
(6) *Agnes knew that I was hot and tired.*
(7) *Agnes said that I was drunk.*
(8) *The policeman believed what Agnes said.*

Attitude verbs combine with sentences to produce verb phrases which are used to classify individuals. Notice, however, that the interpretation of a particular report involving any of (4) to (8) features the interpretation of the embedded sentence, *not* its meaning. A rather straightforward semantic approach, and a first approximation of our own, is to take the attitude verbs as expressing a relation between an individual and the interpretation of the embedded statement. An utterance of (6), for example, would express a relation between Agnes and the fact that I was hot and tired, a complex involving me, two properties, and a location.

This approach to the attitudes exemplifies what Donald Davidson calls "semantic innocence":

> If we could but recover our pre-Fregean semantic innocence, I think it would be plainly incredible that the words "the earth moves," uttered after the words "Galileo said that," mean anything different, or refer to anything else, than is their wont when they come in other environments.[4]

[3] For a discussion of some of these issues, see the second author's "Perception, Action and the Structure of Believing," forthcoming in a Festschrift for Paul Grice edited by Richard Grandy and Richard Warner.

[4] "On Saying That," reprinted in Donald Davidson and Gilbert Harman, eds., *The Logic of Grammar* (Encino, Calif.: Dickenson, 1975), p. 152. Originally published in *Synthese*, XIX(1968/69): 130–146.

On the approach just sketched, the embedded statements and their constituents have exactly the same meaning and interpretation as when they are not embedded.

Traditional objections to the innocent approach have been based on the belief that the only plausible interpretation (reference) of a sentence is its truth value, which obliterates the subject matter of the sentence. The belief that a truth value is the only plausible interpretation for a sentence has been supported by a formal argument which we call "the slingshot." We have shown elsewhere that this argument depends on ignoring from the start the possibility of a situation-based semantics.[5]

Once this objection has been removed, the innocent approach seems quite natural. By focusing on the interpretation (not meaning) of the embedded sentence, it allows us to account for the perspectival relativity of the embedded sentence in a straightforward way.

In the case of first-person, present-tense reports of attitudes, an expressive sentence is used as the embedded sentence in the report. I report the belief I would express with *I am sitting* with *I believe that I am sitting*. But, in general, the sentences we use to report another's attitudes, or they themselves use to report their own past attitudes are not the sentences they would use or would have used to express those attitudes. Thus you will report my belief with *He believed that he was sitting*, not *He believed that I am sitting*. The same point carries over to the other attitude verbs (AV's). Attitudes are attitudes toward propositions. An attitude report NP AVα gets at a proposition P by using an embedded sentence α whose interpretation from the *speaker's* perspective (d, c) is P. The agent a $(= {}_{d,c}[\![\text{NP}]\!])$ would have to use some expressive sentence α' whose interpretation relative to his own perspective (d', c') would also be P:

$$P = {}_{d,c}[\![\alpha]\!] = {}_{d'c'}[\![\alpha']\!]$$

What might be called the "received theory" of the attitudes (setting aside *see* and other perception verbs) goes like this. Attitudes are relations toward sentences, sentence meanings, senses of sentences, or mental representations which are taken to be something like sentence meanings. An attitude report NP AVα reports the agent's attitude toward the sentence α or toward a mental representation somehow associated with α. This is a *"de dicto"* attitude report. The sentence is not used innocently to refer to what it usually re-

[5] "Semantic Innocence and Uncompromising Situations," *Midwest Studies in Philosophy*, VI (1981): 387–403.

fers to, but to refer to itself, its meaning or sense, or to a mental representation.

To account for the phenomenon we have been discussing—the disparity between the speaker's embedded sentence and the agent's expressive sentence, the received theory admits that attitudes are sometimes *reported* in a different way, but maintains that the attitude itself is an attitude toward the received sort of object. Thus, in so-called *"de re"* reports, some parts of the embedded sentence are used not to contribute their meaning (sense, etc.) but to identify, say, an individual *b*. Such a *de re* report, it is claimed, means that the agent has the attitude toward a sentence or meaning that has *b* as the reference of one of its parts. Problems with tense are usually ignored, but would presumably be handled in a similar manner.

There are serious problems with all versions of the received theory—these problems being our impetus for working out an innocent semantics. In the first place what seems to us to be a straightforward phenomenon gets, on the received theory, an extremely complicated explanation that has never been worked out in detail [consider (8), for example].

The idea that attitudes are relations toward sentences is plausible in the case of saying, scarcely plausible in the cases of belief and knowledge, and just wild in the case of perception. And, even in the case of saying, the theory does not run at all smoothly—as Davidson makes clear.

When we turn from sentences to meanings, senses, or mental representations different problems beset us. Frege's notion of sense is often appealed to as if it were a well-developed technical tool. But it is not. Attempts to work out a full-fledged theory of senses meet with serious technical problems, problems that reflect philosophical objections to the very notion of sense.

The index or "possible worlds" semantics developed for modal logic, as adapted for the attitudes, offers us yet another alternative designation for the embedded sentence—its "intension," the set of possible worlds where the sentence is true. Even if one thinks that the primitive idea of a possible world makes sense, the problem of logical equivalence arises. Consider for example,

(9) Fred sees Betty enter.
(10) Fred sees Betty enter and (Sally smoke or Sally not smoke).

We certainly cannot go from (9) to (10), however logically gifted Fred may be. If we did, we should have to admit that Fred either saw Sally smoke or saw Sally not smoke, even though he has never laid eyes on Sally. The admission would be forced by the principles:

If Fred sees P and Q, then Fred sees Q.
If Fred sees P or Q, then Fred sees P or Fred sees Q.

(We wouldn't expect omnipercipience, even among the logically omniscient.[6])

Situation semantics, and semantic innocence, resolve the problem of logical equivalence. Logically equivalent sentences, even in the same discourse situation, are not assigned the same proposition; different subject matters give different sets of situation types. This is the dividend of using partial functions freely in the development of situation semantics. Indeed, from the point of view of situation semantics, the phrase *logically equivalent* should be used for sentences true in the same situation types or courses of events, not for those which satisfy the weaker condition of being true in the same total types or courses of events. The phrase *logical equivalence* having an entrenched use, however, we refer to this stricter relation as *strong equivalence.*

Our innocent approach, then, is straightforward, natural, solves some problems, and avoids others. There are, however, some difficulties.

INNOCENCE THREATENED

In this section we list four problems that threaten our account of the attitudes, problems that point to a missing constituent in our theory.

The logic of the attitudes. There are a number of facts involving the attitudes which seem to require a semantic explanation. These are especially clear in the case of epistemically neutral perception reports [*sees* versus *sees that* as in (4) above].[7] We have stated two of them earlier:

(i) if a sees ϕ and ψ, then a sees ϕ and a sees ψ.
(ii) if a sees ϕ or ψ, then a sees ϕ or a sees ψ.
(iii) if a sees ϕ, then ϕ.
(iv) if a sees $\phi(t_1)$ and $t_1 = t_2$, then a sees $\phi(t_2)$.

Our account so far provides an explanation only for (iv).[8]

Opacity. The astute reader will have noticed that our account violates Frege's and Russell's beginning wisdom on the attitudes—the

[6] See the first author's "Scenes and Other Situations," this JOURNAL, LXXVIII, 7 (July 1981): 369-397.

[7] See Barwise, *op. cit.* for more detail on these points.

[8] Space precludes a discussion of logical relationships between the attitudes, such as the claim that knowing involves believing. We hope the reader will be able to see more or less what we would say, from our discussion of the individual attitudes.

claim that substitution of co-referential expressions does not preserve truth value in attitude reports. After all, if our account predicts (iv) above, it is going to make a similar claim for all the attitudes. By taking the attitudes to be relations to real objects, properties, and relations, we are committed to the claim that they are, in some sense, transparent.

Missing objects of the attitudes. Our account has taken the objects of the attitudes to be propositions. In connection with seeing and saying, this is philosophically unsatisfactory. With seeing it misses the connection with what was actually seen, and with saying it misses the connection with what was actually uttered. Suppose that speaker a, in a discourse situation d and with connections c, says

(11) *b said that* ϕ.

We have seen that ϕ itself cannot in general serve as b's actual utterance. But surely it follows from the truth of (11) that b actually uttered *something*, that there is some sentence ψ uttered by b such that, from b's discourse location d' and with b's connections c', ${}_{d,c}[\![\phi]\!] = {}_{d',c'}[\![\psi]\!]$ and $c(\text{said}) = l_{d'}$. Similarly, if a truly said

(12) *b saw* ϕ.

then what b actually saw was a scene where ${}_{d,c}[\![\phi]\!]$ was realized, not a proposition.

Cognitive content of the attitudes. Missing in our account is the fact that the attitudes have something to do with minds (or brains) and cognition. Just as saying requires the agent to utter something meaningful, and seeing requires the agent to see something with his eyes (a part of the brain), so too believing and knowing require the agent to be in a meaningful cognitive state. Part of what attitude reports give us is information about the agent's cognitive state. That's what makes attitude reports useful in explaining and predicting what people will do. People with similar perceptions, beliefs, and desires behave similarly.

Concentrating on this aspect of the problem makes Fregean senses seem attractive. By interpreting an attitude as a relation to a sense, or "mental representation," as some versions of the theory would have it, one can see the object of the attitude as classifying cognitive states. On this theory, similarity of attitudes points to similarity of states, apparently explaining similarity of actions.[9]

[9] But see Perry, *op. cit.*, where it is shown that this theory doesn't work.

On our theory, however, different people could believe exactly the same thing in countless different ways. The theory so far does not reflect any cognitive similarity at all.

These four problems are interrelated, and all require us to ponder just what it is we are doing when we attempt a semantic theory of the attitude verbs, or of any other "nonlogical" words for that matter. And it is there our defense rests.

INNOCENCE DEFENDED

If simple statements describe situations, then attitude reports must describe situations involving the attitudes, perceptual situations **p** in the cases of *sees* and *sees that,* epistemic situations **k** in the case of *knows that,* doxastic situations **b** with *believes that,* and utterances **u** in the case of *says that.* But just what is it that we are saying about a situation when we say that in it *a sees that* ϕ or *b says that* ψ? Just what is it about the agent that is missing in our earlier account? And what is it about these attitude situations that makes them classifiable with embedded sentences and, hence, according to innocent semantics, with realistic propositions, propositions not in general true of the attitude situation? To answer these questions we must make a slight digression.

Structural Constraints. Things cannot fall out just any old way. There are all kinds of constraints on the types of situation that can actually arise and on the course events can actually take. Some constraints arise from rather obvious properties of and relations between relations. (Kissing involves touching, being a grandfather involves being a father.) Others arise from natural laws. Still others are rather temporary and somewhat accidental (typing used to involve making keys move). A native speaker of a language normally understands many of these constraints and uses this knowledge in discourse. The felicity of exchanges like the following can only be explained relative to such constraints.

> "Did you kiss me?" "I didn't touch you."
> "Is it hot out?" "Well, it's snowing."
> "Why aren't you typing?" "The keys are stuck."

Traditional semantic theories, recognizing the importance of such constraints, attempt to impose them via "meaning postulates" on expressions of language. This strikes us as just backwards. We believe these constraints on courses of events are (except in the most singular cases) independent of which natural-language expressions (if any) designate the constituent objects, relations, and locations.

When the relevant constituents are clearly individuated, it is possible to represent the constraints fairly clearly:

If σ_l(kiss, a, b) = 1 then σ_l(touch, a, b) = 1
If σ_l(bachelor, a) = 1 then σ_l(married, a) = 0
If σ_l(kick, a, b) = 1 and l O l' then $\sigma_{l'}$ (kick, a, b) $\neq$ 0
If σ_l(snowing) = 1 then σ_l(hot) $\neq$ 1

In other cases, it may be rather difficult. For example, it would be impossible to spell out all the constraints on σ^* imposed by σ_l^*(walk, a) = 1.

Systems of constraints can be used for a variety of purposes. A course of events σ is *structurally complete* relative to a set C of such constraints if σ satisfies each constraint in C. A course of events σ is *structurally coherent* with respect to C if σ is part of some complete σ'.[10] If it is not coherent, it is *incoherent.*

A constraint is *correct* if σ^*, the actual course of events, satisfies the constraint. A set C of constraints is correct if each constraint in C is correct, that is, if C correctly captures constraints on the way things can actually happen so that σ^* is complete with respect to C. If C is correct, then every part of σ^* is coherent with respect to C. No part of the actual course of events can be structurally incoherent, though it might be structurally incomplete.

If a_d is an organism in the world, its biological endowment and what it has learned from past experience will lead it to act in accord with certain correct constraints—to be *attuned* to these constraints. As we have seen above, people are attuned to all kinds of constraints they cannot actually state—like all the things that are involved in walking. This is not surprising, since fish are attuned to certain natural laws of water and swimming, and they can't say a word.

But when we are doing the semantics of some word like *kiss* or *walk,* we are forced to reflect on the constraints on kissing and walking with which native speakers of English are attuned and which are reflected and exploited in their use of English.

The same applies to the attitude verbs we have been considering here. Structural constraints come in with attitude reports in two ways. First, there are all kinds of correct structural constraints on attitude situations, just as there are on kissings and walks. Secondly, though, there are also correct structural constraints with which the agent of an attitude situation is attuned and which effect his atti-

[10] σ_1 is *part* of σ_2 if dom(σ_1) $\subseteq$ dom(σ_2) and, for each l ε dom(σ_1), $\sigma_1(l) \subseteq \sigma_2(l)$.

tudes. We are interested primarily in spelling out constraints of the first type, leaving the second to other parts of science.

When we concentrate on the constraints of the first sort which are clearly reflected in language, we find a striking difference between factives (*sees, sees that, knows that*) and nonfactives (*believes that, says that*). The difference shows up most clearly at the extremes with the epistemically neutral *sees* (*Bill saw June win*) and with *says that* (*Bill said that Jane won*); so we begin by discussing and then comparing these two.

Seeing. There is a variety of uniformities across visual perception situations. One sort is built directly into the structure of perceptual reports, but other uniformities are needed to explain the uses we can make of perceptual reports. With epistemically neutral *see* statements, we treated sees as a relation between an agent a and an unlocated proposition P. This is the way language works: *a sees* ϕ focuses on a, on seeing, and on *what is true* of what a sees, $P = {}_{d,c}[\![\phi]\!]$. But the semantic properties of such sentences listed in the first problem reflect a different uniformity—namely, the scene that a visually apprehends. Seeing involves a visually apprehended scene. A scene is an actual situation $\langle l, s\rangle$, but its type does not include everything that happens at l, only that part which is visible under the relevant conditions. These conditions include the direction and distance of the agent from l, the lighting conditions, and much besides. In terms of scenes, we can state the following constraint:

$\sigma_l(\text{sees}, a, P) = 1$
iff there is a scene $\mathbf{s} = \langle l, s\rangle$ such that $\sigma_l(\text{sees}, a, \mathbf{s}) = 1$ and $s \in P$.

All the semantic principles involving *sees* listed in the previous section fall out of this structural constraint. The constraint draws out another uniformity in visual situations, the visually apprehended scene. In doing so, it gives us an alternative indirect way of classifying individuals, by what they saw. That is why we can say

Mary saw a truck stop in front of her. Bill saw it too.

It seems that epistemically neutral reports of visual situations report primarily on what is true of a visually apprehended scene. It is not hard to imagine why language should give us a mechanism for such reports. One need only think of scouts, whose job is to scan the horizons for signs of hostile pioneers. Here we use the perceptual report as evidence, about what the inspected world is like, be-

cause the inspected world, not the agent, is what we are really interested in.

But we also use perceptual reports to *explain* the activities of agents as, for example, when we say that Mary hit the brakes because she saw a truck stop in front of her.

To explain and predict activities of agents, we need to find principles of classification which are projectible onto activities, that is, similarities among agents which lead to similar actions. Given the complexity of the causes of action, these connections will not be simple. But the idea is that any adequate or near-adequate theory—as the theory that supports our explanation of why Mary hit the brakes surely is—must work with a supply of states of the agent which are systematically related to other states and ultimately to activities.

Now we can see the impact of relativity and efficiency. *The chosen uniformities do not by themselves supply all the states or principles of classification we use.* Let us consider our explanation of Mary's hitting the brakes. Consider the class of perceptual situations determined by *saw a truck stopping in front of her.* Clearly, there are many relevant differences. One who sees a truck stopping a mile away will not hit the brakes, nor one who sees a truck stopping in the far lane.

In these last two examples, we have narrowed the classification, in two different ways. In the second case, we augmented the embedded sentence; in the first we considered the distance between the agent and parts of the scene. The general picture that emerges is this. The chosen uniformity—truth of a given proposition in the visually apprehended scene—is not a principle of classification which supports explanation by itself. But it is a part of such a system. The chosen uniformity, together with other factors, gives us a system of (abstract) states useful in prediction and explanation of the agent's activities. Even when we explain by reporting an attitude, we rely on an understanding of the other factors. Thus, in the above explanation of Mary hitting the brakes, the listener limits the other factors in such a way as to make the explanation work—i.e., assumes that the truck was in front of Mary, and not very far ahead.

Says. In seeing, the visually apprehended actual situation plays a crucial role in the classificatory scheme. With a nonfactive, like *says,* there need be no actual situation to support the classification. When we say, *Bill said that Jane won, Jane won* is not serving to classify some actual situation to which George has some relation, say "assertive apprehension." How then does the classification work?

The answer is easy to see or hear. We use *says* to classify utterances. The uniformities across utterances are the very uniformities that we have developed an account of above. Utterances involve discourse situations, connections, and expressions. The chosen uniformity is the interpretation of the utterance.

Indeed, we use the word *says* in two different ways, one that concentrates on what is said in the sense of interpretation, the other in terms of the words uttered. For the latter we use *says* with quotation marks around the embedded sentence. These two uses of *says* focus on two ways of classifying utterance situations. One focuses on the uniformity of interpretation, the other on the uniformity of the meaningful sentence.

Mary said that I was in danger.
Mary said "You are in danger."
Mary said "He is in danger."
Mary said "Watch out!"

Notice that neither of these uniformities can be uniquely determined by the other. The proposition that is stated is absolute:

$\{\sigma \mid \sigma_I(\text{in danger}, a_d) = 1\}$

The expressions *You are in danger* and *He is in danger* are not. The efficiency and relativity of language make it impossible to get from either of these to the other in a unique way.

Note that the two different ways of classifying agents, provided by the two different senses of *say*, provide very different classes of agents, classes that are relevant to different sorts of generalizations. Suppose Hugh says *I am a killer*. Then he belongs to two different but overlapping classes, those who say *I am a killer* and those who say that Hugh is a killer.

Even though these uniformities do not uniquely determine one another, given additional information one sort of classification can lead us, more or less smoothly, to the other. To say *Hugh said that he was a killer* does not *automatically* classify him as an utterer of *I am a killer*, but it suggests it very strongly, since this is the normal way for Hugh to say that he is.

These implications are involved in the explanations of actions by reference to "what was said." Consider for example,

Bill jumped out of the way because he heard Mary say that he was in danger.

As an explanation this makes sense only if Mary said that Bill was in danger in a way that conveyed to him a sense of danger—if she used some expression that is uniform across situations where

the addressee is in danger. *Watch out!* and *You're in danger* are such expressions. *He is in danger* won't, in general, do. What explains Bill's behavior is the existence of a way of saying that he is in danger that is systematically related to situations where the addressee is in danger. It is not just the relativity of language that matters, the fact that there are lots of ways of saying the same thing, but the efficiency of language. An expression like *Watch out!* can be used in many situations to warn of danger.

To complete our theory of saying, then, we need merely to exploit situation semantics (and the other use of *says*) in stating a structural constraint:

If σ_l(says that, b, P) = 1,
then there is an utterance $u = \langle \psi, d', c' \rangle$ such that σ_l(says, d', c', ψ) = 1, where $b = a_{d'}$, $l = l_{d'}$, and $_{d',c'}[\![\psi]\!] = P$.

Now let us compare seeing and saying. The rationale behind the chosen uniformity is quite different. In seeing, the realistic proposition directly classifies an actual situation, and so indirectly classifies the agent who visually apprehends the situation. (Note that we have a theory of *direct* perception, and indirect classification of perceivers.) But, in *saying*, the proposition cannot work their way, for there may be no actual situation that the proposition fits.

For this second scheme to work, there must be something that "fills the gap" left by the absence of any classifiable actual situation. This something is the uttered sentence, or, more plausibly, the utterance of a meaningful sentence with a certain set of intentions. Instead of a relation to an actual situation that the proposition characterizes, we have a relation to a meaningful entity which, in the utterance, has the proposition as its interpretation.

However, there is an important difference between the relation to the actual situation and the role of the uttered sentence.

In seeing, the proposition connects to the agent "through" the apprehended scene. In saying, the proposition connects to the agent through the produced sentence. But the proposition is true in the situation of which the scene is a part, quite independent of the agent's location, connections, history, etc. The situation, we might say, gives us a pool of propositions; the other factors merely influence the way the agent can apprehend the scene of which it is true.

But, in saying, the proposition does not in general (if ever) connect to the isolated meaningful sentence, but only to the entire utterance. The proposition does not serve to classify *one* of the factors, and thereby the whole, but characterizes the whole by the

relationship among the factors. (A special case would be the utterance of a sentence whose meaning uniquely determined its interpretation, if there are any.)

When we move to the question of the interest in the chosen uniformity across utterances, we find a similarity with seeing. *Says that* is designed to tell us what the world is like if what the agent says is true. But, as with seeing, other uniformities are crucial when we use *says* to explain and predict activities of the agent, or of those who hear or read the utterance.

Sees that, knows that. Suppose the identical twins June and Jane entered the marathon and Bill saw that one of them won. In fact it was June, but Bill can't tell them apart. If asked which one, he couldn't say. Cases like this bring out the difference between non-epistemic *sees* and *sees that* and *knows that.* If Bill saw one of the twins win, and June won, then Bill saw June win. But even though he saw that one of them won, and knows that one of them won, he doesn't see that June won, or know that June won.

Seeing that involves scenes (or, more generally, courses of events), but the relation is less direct than in the case of *sees.* The structural constraints are (roughly):

σ_I(sees that, a, P) = 1
iff there is an event σ_0 (possibly a scene s) such that
(i) σ_I(sees, a, σ_0) = 1
(ii) there is a system C of correct structural constraints with which a is visually attuned such that every σ containing σ_0 which is complete with respect to C is in P.

In the above example there is a certain visual property p such that Bill is attuned to:

If $\sigma_I(p, x) = 1$, then $\sigma_I(p_{\text{June}}, x) = 1$ or $\sigma_I(p_{\text{Jane}}, x) = 1$.

where p_{June} is the property of being named June. Any structurally complete σ containing the event Bill saw will have $\sigma_I(p, x) = 1$ and $\sigma_I(p_{\text{June}}, x) = 1$, but this σ_0 was not complete.

The difference between reports using *sees* and reports using *sees that* lies partly in these constraints, but also on different interpretation strategies for noun phrases and verb phrases in the embedded sentences. We discuss this briefly as "value loading" in the next section.

One can give a similar structural constraint for *knows.* The basic idea is that to know is to be attuned. We will be simple-minded and pretend that vision is the only form of perception, for exposi-

tory purposes. Then one could say

σ_l(knows that, a, P) = 1
iff there is a course of events σ_0 such that
(i) σ_l(sees, a, σ_0) = 1
(ii) if C' is the set of all correct structural constraints with which a is attuned, then every σ containing σ_0 which is complete with respect to C' is in P.

The reader will notice that the only difference here between *sees that* and *knows that* lies in the fact that a wider set C' of correct constraints is admitted for *knows that*. This presumably has something to do with the tendency to say we see that something is the case when we mean we know it.

As with seeing, the treatment of *sees that* and *knows that* explains the many sorts of uniformities, other than the "chosen ones," across epistemic situations. In particular, the structure of knowledge allows for knowing the same thing on the basis of different sets of past experiences, a difference that might be relevant to how the knowledge affects one.

Believing. It is usually thought that knowing that P implies believing that P. And in language we usually assume that, if the speaker says that P, he believes that P. In this regard, believing seems more or less intermediate between knowing and saying. But, from a realist perspective, belief is by far the most puzzling of our four attitudes. For where or what is the real invariant in various actual doxastic situations which supports their classification together as situations in which an agent believes that P (where P is a realistic proposition)?

With seeing, seeing that, and knowing that there is an actual course of events of which P is true. With belief there need be no such; so believing seems more like saying that. But in the case of saying that there is the expression actually uttered, something real, which, together with the discourse situation and speaker connections, gave rise to the proposition. What is analogous is the case of belief?

It seems that the realist, if he believes in belief, is forced either into a metaphysics that includes real but not actual situations (an outlandish move that surely no one would advocate) or into a metaphysics that countenances real "belief states," some kind of abstract but real invariants across actual doxastic situations, invariants that support their classification by realistic propositions the way sentences support propositions in the case of saying.

This is where one might think that something akin to Fregean

senses comes in, "complete and eternal thoughts" grasped by minds. We could modify the Fregean account so that the reference of a thought T was a realistic proposition $P = \text{ref}(T)$ and use the structural constraint:

σ_1(believes a, P) = 1
iff there is a thought T such that σ_1(doxastically grasps, a, T) and ref(T) = P.

This would allow us to capture the relativity of belief, the fact that different people can believe the same thing in different ways, by having different thoughts. But this would be a serious mistake! It is just as important for belief states to be both efficient and relative as it is for sentences. There are other factors that play a role in going from the state S_i to the proposition: $P = F(S_i, \ldots ?)$. What are these other factors?

The agent, of course, is one such factor. When it is in what we might call the "I'm in danger" belief-state, its beliefs are about itself and its present location in time and space. And there is no reason to suppose that other properties of the agent, say his height or education, might not play a role in the interpretation of his belief state. We lump all this into an agent situation $d = \langle s_d, a_d \rangle$ where $s_d = \langle l_d, s_d \rangle$, a_d being the agent, l_d its location, and s_d the facts about a which are needed for interpretation. This d is analogous to the discourse situation in the case of saying.

But of course we can have beliefs about things other than ourselves and our present location. We have connections with objects, relations, and locations which arise through perception, and these connections help determine what our beliefs are about.

Thus talk about belief presupposes an abstract classification system $S_1, S_2, \ldots$ of *states* and a relation bel that holds between states, agent situations, connections, and courses of events:

bel(S_i, d, c, σ)

Then we impose the constraint that σ_1(believes that, a, P) = 1 iff there is a d, c, and S_i such that $l = l_d$, $a = a_d$, $\sigma_1(S_i, d, c) = 1$, and $P = \{\sigma \mid \text{bel}(S_i, d, c, \sigma)\}$.

The relation bel is analogous to [[ ]]. Just as [[ ]] identifies a proposition relative to an expression, discourse situation, and connections, bel identifies a proposition relative to a belief state, doxastic situation, and connections: the proposition that a person in that state, in such a doxastic situation, with such connections, believes. The postulation of such a relation and such a system of "meaningful" states is presupposed by the way we use *believes*. This ap-

proach to belief seems to us to fit well with a number of approaches to the philosophy of mind, which emphasize how attributions of mental states are connected with activities of species and of individuals.

In the case of agents that speak a language L, it is very tempting to assume that the meaningful sentences of L can be embedded in the structure of the belief states, that there is a function $S(\phi)$ from sentences of L into the system of belief states. This suggests that belief states have a certain "syntax" analogous to the syntax of L and that this "syntax" is important in the analysis of bel, just as language syntax is important in the analysis of ${}_{d,c}[\![\phi]\!] = P$. This temptation should be distinguished from something we do not find very tempting: the view that believing consists in having some relation to some sentences of some language.

OPACITY

The interaction of attitude verbs with singular terms (proper names and definite descriptions, for example) was a driving force behind the theories of Russell and Frege—different as these theories were. It is certainly possible for George IV to wonder whether Scott is the author of Waverly without wondering whether Scott is Scott, and it is possible for one to believe that the morning star is a planet without believing that the evening star is a planet. These facts caused Frege to say that, within the scope of an attitude verb, an expression refers to its usual "sense," not to its ordinary reference. Russell introduced "logical form" and argued that definite descriptions do not denote but rather contribute the defining properties to the proposition properly understood. (Many contemporary theories appeal to both sense and logical form.)

We do not appeal to either sense or logical form, but handle these problems basically with the resources already at hand. There is not space to explain our treatment of names, although the reader can probably guess how such notions as inverse interpretations and connections allow us to replace such old questions as "Do proper names have sense?" and "What are the truth conditions of a sentence with proper names?" with more tractable questions. We shall explain the basic ideas behind our treatment of descriptions, a treatment that has many Russellian features, but does not threaten us with atomism or require appeal to logical form.

To simplify discussion we ignore issues of time and place, so that we can deal with situation types rather than courses of events. We also restrict ourselves to descriptions α that are not sensitive to discourse situation d or connection c, so that we can write $[\![\alpha]\!]$ for the interpretation ${}_{d,c}[\![\alpha]\!]$, again just to simplify discussion.

The interpretation of a definite description is a relation between situation types s and individuals a:

$$[\![\textit{The } \beta]\!](s, a) \text{ iff } \{a\} = \{x | [\![\beta]\!](s, x)\}$$

This relation can also be viewed as a *partial* function from situation types s to individuals a. Using standard function-argument notation, we can write $a = [\![\textit{The } \beta]\!](s)$. This function sets up a mutual constraint between s and a. Given an s in its domain, we can use *the* β to refer to $a = [\![\textit{the } \beta]\!](s)$. Or, given an a, we can use *the* β to claim that the situation s is one where $[\![\textit{the } \beta]\!](s) = a$. Or it can be used simply to say that, whatever s and a are, $[\![\textit{the } \beta]\!](s) = a$.

EXAMPLES. (i) I walk into Alfred's study where he sits with his dog Clarissa. He says "Be careful. The dog has fleas." The situation s_0 we are in makes it clear that he is referring to Clarissa (= $[\![\textit{the dog}]\!](s_0)$).

He has asserted the realistic proposition:

$$\{s_1 | s_1(\text{has fleas, Clarissa}) = 1\}$$

Notice that, if I believe him, then what I believe is not that there exists a unique dog that has fleas, but rather that this particular dog has fleas.

(ii) Now there are several dogs in the room. Pointing at Clarissa, Alfred says *This is the dog that bites.* Here the definite description *the dog that bites* is not being used to pick out Clarissa, but rather to attribute to her the property of being the unique dog that bites. The proposition is

$$\{s_1 | \text{Clarissa} = [\![\textit{the dog that bites}]\!](s_1)\}$$

(iii) Now we are in a situation where Agnes once told me of a certain individual a, *She is a fool.* Agnes is a shrewd judge of character. Accordingly, I counsel you against investing in a's bank by warning you, *Agnes believes the president of First Federal is a fool.* Here the interpretation of my utterance is, essentially,

$$\{s_1 | s_1(\text{believes, Agnes}, P_{s_1}) = 1\}$$

where

$$P_{s_1} = \{s_2 | s_2(\text{fool}, [\![\textit{the president of First Federal}]\!](s_1)) = 1\}$$

The definite description constrains s_1 to contain a unique president of First Federal, a_{s_1}, and asserts that Agnes believes that a_{s_1} is a fool (P_{s_1}).

We use the notation: *a says(believes/knows/sees)that(—(the* $)_j$. . .)

with $j = 0$, 1, or 2 to indicate the readings where $[\![the\ \beta]\!]$ is evaluated at an accessible situation type s_0, as in (i); to constrain the situation type s_1 designated by the whole (as in iii); or to constrain the situation types s_2 described by the embedded sentence, respectively. For $j = 0$ this corresponds to Donnellan's referential use. For $j = 2$ it is his attributive use. The case $j = 1$ is somewhere in between.

These various readings might appear to coincide with different scope readings. But they really reflect a different phenomenon, one that is widely confused with scope. The distinction is most easily seen with indefinite descriptions, like *member of the family*. The interpretation of an indefinite description *a* β (e.g., *a dog, an elephant*) is also a relation between situations and individuals:

$$[\![a\ \beta]\!](s, b) \text{ iff } [\![\beta]\!](s, b)$$

Consider the case where Jack has been murdered. Holmes has assembled all the members of the family and said "One of you has murdered Jack." "What did he say?" asks deaf old Aunt Agnes. "He said that a member of the family murdered Jack," yells Jack's widow Jill.

Now there is certainly nothing wrong with Jill's report, but it cannot be accounted for with only the wide scope/narrow scope distinction. Holmes didn't say of any particular member in the family that he was the murderer, so it is not wide scope. But he didn't say anything at all about family membership, so it is not narrow scope. It is what we would write as *He said that* (*a member of the family*)$_1$ *murdered Jack*. The interpretation is:

$$\{s_1 | s_1(\text{says, Holmes}, P_{s_1}) = 1\}$$

where

$$P_{s_1} = \{s_1 | \text{for some } a \text{ such that } [\![a \textit{ member of the family}]\!](s_1, a), s_2(\text{murdered}, a, \text{Jack}) = 1\}.$$

FOUNDATIONAL ISSUES

Certain foundational issues confront anyone who tries to work out a careful semantic theory of the attitudes. We cannot here discuss the exact guise in which these appear for situation semantics, or the details of our solution. The basic idea is to restrict ourselves to the hereditarily finite set-theoretical objects built out of the objects, relations, and locations at our disposal. Ultimately, this requires us to be more realistic about the sentences, states, and other factors involved in the attitudes. For example, ultimately we define

$$\sigma_1(\text{says}, a, P) = 1$$

by

$$\exists d, e, \psi[\sigma_l(\text{says}, d, c, \psi) = 1 \,\&_{d,c}[\![\psi]\!] = P]$$
$$(\text{with } \sigma_l(\text{says}, a, p) = 0, \text{ otherwise}).$$

This allows us to avoid having propositions as arguments of situation types. At this point, our theory has led us up a spiral. We started from a realism toward situations in the world, were forced to be realists about objects, properties, relations, and locations. This forced upon us a philosophical realism toward cognitive states and activities. In the end, this allows a slight abandonment of pure innocence in favor of a sort of worldly innocence, which we hope that some readers may find attractive.

JON BARWISE AND JOHN PERRY
University of Wisconsin/Madison and Stanford University

DEEP INTERPRETATION*

But if you were to hide the world in the world
so that nothing could get away, this would be
the final reality of the constancy of things.

Chuang Tzu (tr. Burton Watson)

THERE is a concept of interpretation abroad these days which, though it arises in particular connection with texts, has little to do with matters that call for interpretation in the rather more routine acceptance of the term: with whether or not a certain ambiguity or inconsistency is intended and, if inadvertent, with how such flaws are to be resolved—with how the text is to be *read*. Thus the chronology appended by Faulkner to the text of *Absalom, Absalom* happens to be inconsistent with the chronicle one may recover from the notoriously tortured narrative of the novel, and there is an initial question of whether Faulkner got it wrong or whether the text of the novel is to be amended, or if it was deliberately planted to excite an even deeper reflection on time, voice, and narration than the already complex narrative structure alone would arouse in the literary consciousness of the reader. More important, the chronology must on

* To be presented in an APA symposium of the same title, December 29, 1981. David Hoy will comment; his paper is not available at this time.

0022-362X/81/7811/0691$01.60