

$$Mp \equiv \bigwedge K_i p \wedge K_i Mp.$$

According to this definition, everyone knows p and everyone knows it is mutually known. Under normal communicative circumstances it is reasonable to commit to a belief in mutual knowledge of p , which can be done in finite time. Three things make one suspicious of S&W's rejection of mutual knowledge: First, coded communication does, of course, occur; second, it is later admitted that communication need not, indeed cannot, be guaranteed to succeed, even though this requirement was imposed on coding; and third, since manifestness is only a dispositional property, in S&W's definition, it is hard to see how mutual manifestness allows the essential inferences to be made about the actual cognitive states of others. The concept seems to suffer the drawbacks of Hintikka's (1962) deductive-closure model of knowledge without the benefits of an explicit model such as that of Konolige (1986).

Fortunately, the detour via mutual manifestness can be avoided without undue strain on the rest of the theory of relevance. Perhaps a better reason to reject the code model is simply that the distortions caused by reformulating the process of contextual inference as a sign-message correspondence system render the model practically useless as a generator of insights into complex communication phenomena.

S&W propose two main constructs to flesh out their rational approach: a model of human inference and a definition of relevance, which plays the role of the human cognitive goals in the rational communication process. These cognitive goals are assumed, in some "trivial, speculative remarks," to be the enlargement of the set of explicitly available facts known and the minimization of processing effort. It might be more consistent with a rationalist approach to include the hearer's goals as a constraint on which new facts are inferred, as in much current work in artificial intelligence.

A model of an inference mechanism is required to give substance to claims about which facts become known as a result of communication, and with how much effort. To a reader familiar with artificial intelligence and logic, the proposals for this mechanism put forward in *Relevance* seem somewhat strange. Here are the problems associated with several claims:

Claim: Inference consists of unfettered forward chaining to termination. This is absurdly impractical, as actual experience demonstrates (Schank 1979). Control methods must be used to ensure that only useful inference paths are explored.

Claim: All inference rules are explicitly represented. Such a scheme cannot succeed, because whatever combines the inference rules with their premises must also be an inference rule, but procedurally represented; moreover, almost all actual inference machines incorporate their inference rules by means of procedures because this is far more efficient, and inference rules never change.

Claim: All inference is eliminative. This would disallow the use of commutativity laws and axioms defining inductive domains, for example.

Claim: Assumption schemas are not complete propositions, yet can be used in inference. A formal object without a truth value cannot, for example, contradict anything. As admitted by S&W (chap. 2, n. 10) these schemas are really just existential propositions.

Claim: Assumption strengths, or degrees of confirmation, are not explicitly computed, but are merely produced as the result of nonlogical processing of assumptions; hence confirmation theory is irrelevant. This seems to be a confusion between the specification of a logical confirmation process and its implementation. By the same token, one could say that a theorem prover using inference procedures rather than explicit inference rules is not a logical system. S&W's rejection of a computational approach to confirmation is especially odd in light of the fairly standard (although shaky in places) probabilistic inference procedure they advocate.

Fortunately, again, these deficiencies can be repaired without undue damage to the rest of *Relevance*. However, it seems odd that little or no reference is made to the considerable work in artificial intelligence on models of plausible inference, both for problem-solving in general and for language understanding in particular. (Kanal and Lemmer [1986] and Norvig [1987] contain worthwhile discussions.) The procedures S&W propose for memory access and context construction are woefully underspecified.

The two principles (it seems there really are two) of relevance and least effort are put to work to explain a large number of comprehension phenomena. It is here that the work cries out for a formal implementation, because in any number of cases we are expected to sit idly by while yet another *ad hoc* rule for estimating effort or for defining and accessing contexts is introduced in order to rationalize a desired interpretation (see, for example, pp. 187, 200, 219–20, 221). Effort, in particular, is far too flexibly defined to be a theoretically adequate tool for a real pragmatic system, and there seems to be no way to obtain more than introspective evidence. Although I will leave it to the linguists to assault with deadly example-sentences, it seems unlikely that the minuscule amounts of effort required by S&W's theory, their motivating picture of mental exhaustion notwithstanding, would play a significant role except in highly time-pressured situations. To place all the burden of selecting correct interpretations on an effort principle, implemented as an ordering on context accessibility, abdicates the responsibility for a theory, passing it on to those studying memory organization, who are already overloaded with ordering constraints passed on by other theories such as the possible-worlds theory of counterfactuals.

In general, one gets an uneasy feeling about the value of some of the explanations proposed in *Relevance*, akin to the feeling experienced when reading teleological explanations in evolution or functionalist explanations in sociology. Although, in principle, one might be able to trace all aspects of language back to their roots in cooperative advantage, the effort may be neither practical nor useful. For example, Quine's (1960) model of how words get their meaning is very interesting philosophically, but it should not be part of an explanation of how language is processed. S&W are apparently happy to accept the existence of convention in the area of grammar and the lexicon, although they reject all other uses of convention as *ad hoc*. Conventions of all kinds are, however, self-justifying, since by their existence they enable more effective cooperation. Even if some of the pragmatic processes involved in communication have a rational explanation, it would seem that our first task is to establish exactly what the processes are, along with what conventions they may rely on. The first job of an Englishman after a transatlantic migration is not to find out the rationale behind a society's choosing one side of the road to drive on rather than the other, but to find out which side of the road to drive on.

How relevant?

Pieter A. M. Seuren

Philosophy Institute, Nijmegen University, 6500 HK Nijmegen, Holland

It is gratifying to see Sperber & Wilson (S&W), in their *Précis* of *Relevance*, state openly, at the very end, that they are "well aware that the view developed in *Relevance* and summarized here is very speculative and, as it stands, too general directly to determine specific experimental tests or computer simulations." Nowhere earlier had this been said so explicitly. One is now fully justified in concentrating on the internal logical coherence of the view in question, for if a speculative view has any force, it is mainly on account of its internal structure and logical clarity. S&W's critics, including myself, have always concentrated on precisely these issues, and, as it now appears,

rightly so. In my (very negative) review of *Relevance* (Seuren, 1988) I raise a number of central issues having to do with the coherence and consistency of S&W's speculations, to which Wilson and Sperber (1988b) wrote a (rather piqued) reply. I shall now take up some of the objections left unanswered, or inadequately answered, in that reply.

The central theme in S&W's view of relevance is the fact that "there is a gap between the semantic representations of sentences and the thoughts actually communicated by utterances. This gap is filled not by more coding, but by inference" (Précis, sect. 1.1, para. 3). Thoughts are "conceptual representations" (*Relevance*, p. 2). Among what are considered thoughts, *assumptions* occupy an important place: They are "thoughts treated by the individual as representations of the actual world (as opposed to fictions, desires, or representations of representations)" (p. 2).

The relevance of an assumption in a context is said to depend on the extent of the *contextual effects* in that context, as well as on the effort required to process it in that context: More contextual effects and lesser processing effort lead to greater relevance-in-that-context of the assumption in question. One has to take it, therefore, that an utterance *u*, together with an unknown quantity *i* of inferencing in a context *C* yields a set of thoughts (assumptions) *I* actually communicated by *u*. *I* is graded for relevance according to the wealth of contextual effects and the minimality of the processing effort. This grading is not absolute but relative to other possible assumptions communicated by *u* with other inferencings and/or with other contexts. Each pairing of some *u* with some *i* in some *C* is called a *hypothesis* (p. 165), let us say *H*. Each *H* thus determines some set of assumptions *I*, with an associated relevance value *R*. There is then a *Presumption of optimal relevance*, which makes the hearer select the *I* with the highest *R*, at least in some cases (p. 165): "In some cases, [the task of constructing and selecting a hypothesis] is best carried out by listing all the possible hypotheses, comparing them, and choosing the best one." In other cases, however, "it is better carried out by searching for an initial hypothesis, testing it to see if it meets some criterion, accepting it and stopping there if it does, and otherwise repeating the process by searching for a second hypothesis, and so on." One might now wonder what can be meant by "processing effort," in that the processing of an utterance involves the adequate selection of a context (pp. 137–38): What if *u* yields a very high *R* (lots of contextual effects, at very little processing cost) at hypothesis 358? Would not the processing cost needed to get to H358 outweigh the saving at H358? Wilks (1986, p. 273) observes that this seems to make nonsense of the condition of minimal processing effort. S&W's answer (p. 167) is that, in fact, the interpretation to be selected is "the first one tested and found to be consistent with the principle [of relevance]. . . . The important point is that, given the cognitive environment, given the initial context, and given the stimulus, some hypotheses are more accessible than others, and this means that they require less processing effort." Had the communicator intended the hearer to settle for a less accessible hypothesis, he should have used an utterance that saves the addressee the effort of first accessing more accessible hypotheses (p. 168).

All this raises a number of questions. First, S&W's account requires that, given a context (including a "cognitive environment"), a given utterance *u* determines the order of accessibility of possible interpretative hypotheses. Apart from the empirical problem raised by this requirement (we have promised to let that aspect rest), there is the fact that the function involved is left totally unspecified, crucial though it is in the total framework. Having said (p. 167) that *u*, in virtue of its grammatical and lexical form, gives immediate access to a highly determinate set of concepts, assembling them into a logical form directly usable as an assumption schema, S&W content themselves with stating: "The context provides ways of completing these assumption schemas into full hypotheses." This conclusion is presented as a

fact, and not as a serious white area on a map whose filling in is doubtful. Unfortunately, this type of procedure is typical for *Relevance* (and its Précis): Very often proper analysis and argument have to give way to rhetorical suggestions and demonstration by way of examples. Yet the result is put forward as solid truth. The book would have gained a lot if it had been written less apodictically.

It would also have gained a lot if it had been written more clearly. What one finds is an almost inextricable maze of definitions couched in often unnecessarily opaque and abstruse terminology. One fails to find clear statements of the computational aspects (functions) involved or invoked, or of the rules or principles needed to make the less determinate parts of the system work. This is perhaps best illustrated by holding up for close inspection S&W's notion of contextual effect (CE). We remember that without CEs there is no relevance. What, then, is a CE? It is meant to be an alteration of a given context resulting from the addition of a new utterance (not counting the purely incremental effect of the utterance). The context *C* can be modified in three possible ways (p. 117): Some assumption may be removed from *C*, its "strength" may be modified, or *C*, enriched with *u*, may now possess a richer set of entailments than either *C* or *u* alone. In the last case we have a so-called *contextual implication* (pp. 107–8). Moreover, the strength of any assumption depends on the evidence available for its truth (p. 109); and the highest degree of strength, one infers from p. 113, is certainty. Strength adheres not only to the contents of the speaker's message but also to the listener's recognition of the speaker's intention as part of the set of assumptions transmitted. When a conclusion is drawn, its strength will depend on the relative strengths of the input assumptions (pp. 109–16). But S&W's text is rather verbose but sketchy and imprecise on this point: No precise procedures or formalisms are provided. Moreover, if strengthening helps comprehension (by raising *R*), how can weakening do the same? According to S&W, if *u* is contradictory with some element *e* in *C*, then the weaker of the two will be erased from *C* (p. 115). If *e* is the weaker one and is eliminated, we have a CE, but if *u*, the new utterance, has to be kept out, there is no CE. Erasure of an *e* can be considered the most extreme case of weakening, and this, too, helps comprehension! Would not some further clarification be in order here?

All in all, this centrally important notion in S&W's framework remains without a precise definition. In Wilson & Sperber (1988b) it is claimed that *Relevance* (p. 260, n. 26) does offer a formal definition. In this footnote a "more formal characterisation" is promised. What follows is my rephrasing of the relevant text in more easily accessible terminology and symbolism. Let *C* stand for context, *P* for newly added propositions ("assumptions"); E_C and E_P represent the set of conclusions derivable from *C* and *P*, respectively. The contextualisation of *P* in *C* has no CE just in case (i) $E_C \subset E_{C \cup P}$, and (ii) $E_{C \cup P}$ minus $E_C \subset E_P$. If these two conditions are not both met, there is some CE. But notice that condition (i) is trivially met in all possible cases, since whatever follows from a set of propositions *C* also follows from *C* enriched with any set of propositions. So the only condition is (ii): For there to be a CE there must be a conclusion in $E_{C \cup P}$ which is neither in E_C nor in E_P . But this is precisely the definition given for *contextual implication* (pp. 107–8), and thus, what is defined in note 26 is an unnecessarily complex rephrasing of what had already been given as a definition for *contextual implication*, and does not define *contextual effect*. The elements of "strength" and "erasure through contradiction" are still undefined. It must be added that S&W also say in note 26: "Let two assumptions with the same content but with different strengths count as two different assumptions." But is one to understand that if only strengthening occurs (whatever that may be), then, due to a mysterious change in identity of the strengthened assumption set, there are suddenly also contextual implications which, without the strengthening, would

not have been there? If so, the reader might expect considerably more explanation.

This leads us to S&W's notion of *contextual implication*: A set of assumptions *P* contextually implies an assumption *q* in the context *C* if and only if (i) the union of *P* and *C* nontrivially implies *q* and (ii) neither *P* nor *C* nontrivially implies *q*. This, as we have seen, is identical with the definition of contextual effect (p. 260), except for the specification that the deduction must be "non-trivial." (I take it that "non-trivial implication" is also what is meant in that definition. Otherwise I cannot understand what is meant.) What, then, is "non-trivial implication"? For this we must go to p. 97: A set of assumptions *P* *logically and non-trivially implies* an assumption *q* if and only if, when *P* is the set of initial theses in a derivation involving only elimination rules, *q* belongs to the set of final theses. By "derivation" S&W mean a formal, logically deductive procedure whereby logical consequences (entailments) are computed. S&W maintain, no doubt correctly, that standard first order logic is both too poor and too rich for the proper analysis of natural language entailments: It is too poor because it fails to capture many if not most intuitive entailments, and it is too rich because some formal entailments in logic are counterintuitive. In order to remedy the poverty, S&W invoke a "concept logic," the details of which are not provided. Against the excessive logical power they posit that in language and cognition only elimination rules are used and never introduction rules. This in itself has extremely awkward consequences, since it rules out the conclusion that two people came in if it is indicated that Henry and Jack came in – to mention just one example. Moreover, the notion of elimination rule is left opaque (for details see Seuren, 1988). The nonstandard logic invoked by S&W thus seems not to exist in the sense of being formally available. S&W further claim that, in their logic, the set of entailments associated with any *P* is finite. No evidence is offered for this bold thesis, yet S&W state that "The process [of derivation or deduction] applies to all initial and derived theses until no further deductions are possible" (p. 95).

Let us now turn back to the definition of nontrivial implication as paraphrased above (from pp. 107–8). What is meant by "the set of final theses"? Note that this term is simply used in this definition but nowhere explained or defined. In Seuren (1988) I took it that, in accordance with current terminology in the theory of algorithms, it should refer to the set of ultimate conclusions from which no further deductions are possible. Wilson & Sperber (1988b), however, indicate that this is not the intended interpretation: The term must be taken to refer to the union of the initial and the derived theses (assumptions). Not only does this terminology seem sloppy, but "the set of final theses" is still not defined, and hence neither are the notions of nontrivial implication, contextual implication, and contextual effect.

This is only one of many similar tangles in the book. What, then, remains of S&W's speculative view? Very little, in my opinion. I know of no other attempt of the same broad scope and proportions to explain the complexities of linguistic and non-linguistic comprehension of messages. Perusing both *Relevance* and its *BBS Précis* calls to mind why no alternative approaches have been attempted: The enterprise seems altogether too ambitious, for the time being, to be realistically undertaken. In my view it is much more realistic and fruitful to let the wide and general issue rest for a while, in order to concentrate on detailed empirical analyses of more restricted problem areas. The empirical knowledge thus gathered will prove highly relevant in constraining and shaping speculative hypotheses about the broad kinds of issues that S&W have the courage to tackle. There is plenty of room within the bounds of the empirically accessible for fruitful theory to develop.

The information needed for inference

Carlota S. Smith

Department of Linguistics, University of Texas at Austin, Austin, Tex. 78712

Sperber & Wilson (S&W) argue that verbal communication involves both code and inference. In making inferences people use general abilities and knowledge rather than the specifics of a code; and the fact that people revise and abandon interpretations is strong evidence in favor of S&W's position. Writers often exploit this flexibility, for example, by using unreliable narrators in fiction. Readers must gradually revise their interpretation of a text as they lose faith in the narrator.

The many inferences in S&W's account of communication require several types of information, principles, representations, and computations. The principle of relevance seeks an optimal balance between effort and effect. As an essential first step toward understanding how the principle might actually operate, I wish to consider the information required for inference at the various stages of understanding. The variety makes it somewhat difficult to see how a single principle might apply.

There are three main stages in understanding, according to the programme of S&W: decoding an utterance, arriving at its propositional form, and assessing its import. The last two stages require inference. Consider utterance (b) below, uttered by B in the context of A's uttering (a). A's task is to understand (b).

A: (a) Susan has thalassemia.

B: (b) She is getting married to Bill.

Assume that A, having decoded (b), is in a position to infer its propositional form, or explicatures. The explicatures are based in part on the logical form of (b). To work them out in detail we would need a representation of logical form that could be used in inferences concerning ambiguity resolution, reference assignment, completion or enrichment, and determination of propositional attitude.

Ambiguity: A must decide between the interpretations decoded for an ambiguous utterance, choosing the one most compatible with the situation. Since (b) is not ambiguous, take S&W's example *The child left a straw in the glass*. To resolve the lexical ambiguity the receiver consults encyclopedic memory, which contains stereotypical scenarios that chunk information and are relatively accessible; such a scenario, once found, would give the "drinking tube" interpretation. Memory also contains the "piece of hay" interpretation, and presumably other information about straw as well. But to work out the effort required to consult encyclopedic memory, we need to know how it is organized. Specifically, in what sense is a chunk accessible? What steps are required? What choices are available? In addition, we need to know how situations are represented and how they affect the search of memory. Note that if the sentence above were uttered during a country picnic the second interpretation might be preferred.

Reference assignment: A must determine the reference of the pronoun *she*. Suppose that A finds no third person in the environment and therefore discards the possibility of a deictic interpretation; A then chooses the anaphoric and looks for an antecedent, finding in (a) a noun phrase of the correct gender and number. This information is quite different from that required for ambiguity resolution, as are the calculations involved. To determine the reference of *Bill*, A draws on knowledge of people familiar to both A and B; such knowledge must be represented and computed.

Completion: The present progressive in (b) is vague and must be filled out by inference. It may refer to a marriage in progress or a future marriage. Suppose that A, consulting the immediate environment, fails to find a wedding in progress. But perhaps the wedding is going on somewhere else: A must search memory, perhaps expanding the context, for information on this point.