

They say it's a new sort of engine:
but the SUMP's still there

Karen Sparck Jones

Computer Laboratory, University of Cambridge
Corn Exchange Street, Cambridge CB2 3QG, England
sparckjones%cl.cam.ac.uk@cs.ucl.ac.uk

I shall lump the specific semantic formalisms currently touted together as manifestations of logicism, because the issue is whether the logicist approach to language processing is the right one, not whether one particular formalism is better than another.

The logicist model of language processing is essentially as follows. We use a phrase structure grammar, heavily laced with features, for syntactic parsing; in analysis syntactic processing drives semantic interpretation strictly compositionally, to build a logical form representing the literal meaning of a sentence. This logical form is further processed, both in discourse operations of a larger scale linguistic character, as in (some) pronoun resolution, and, more importantly, in inference on global and local world knowledge, to obtain a filled-out utterance interpretation. Logical form plays a key role, motivated as much by the need for reasoning to complete interpretation as by the need to supply an appropriate input to further reactive problem solving. The use of a logical form to represent the meaning of an input naturally fits the use of a logical formalism to characterise general and specific knowledge of the world to which a discourse refers and in which it occurs.

Whether the logicist position is psychologically plausible is not an issue here: it can be adopted as a base for language processing without a commitment to, say, FOPC as a vehicle for human thought, and indeed can be adopted in a psychologically implausible form, for example with complete syntactic processing followed by semantic processing. The fact that a good deal of development work is being done using a case frame approach is no problem for the logicist either: case frames are either an alternative notation or can only work for the discourse of very restricted applications.

The logicists' problems are those of getting an expressive and tractable enough logic: what sort of logic is powerful enough to capture linguistic constructs, characterise the world, and support common-sense reasoning; and is this logic computationally practical? Determining and representing the knowledge required for a particular system application is not necessarily more of a problem for the logicist than for anyone else. The threat to the logicist's moral, or at least mental, purity is whether a logic which will do the job is really a logic at all. If the world is heterogeneous, our thinking sloppy, and our language uncertain, whatever closely reflects these may be barely worthy of the label "logic".

This is not to rake up the old semantic nets versus predicate logic controversy, or its analogues. We may have a formalism with axioms, rules of inference, and so forth which is quite kosher as far as the manifest criteria for logics go, but which is a logic only in the letter, not the spirit. This is because, to do its job, it has got to absorb

the ad hoc miscellaneity that makes language only approximately systematic. Broadly speaking, it can do this in two ways. It can achieve its results through some proliferation of rules, weakening the idea of inference. Or it can achieve them essentially by following the expert system path, retaining a single rule of inference at the cost of very many specific, individual axioms. It is at least arguable that if the stock of initial propositions is a vast heap of particulars defining idiosyncratic local relationships, the fact that one is technically applying some plain rule of inference to follow a chain of argument is not that impressive: conciseness and generality, which at least some expect a logic to have, are not much in evidence. Precision and clarity may be equally unattainable, at any rate in practice.

I believe the second possibility is already with us, masquerading in the respectable guise of meaning postulates, and that whatever precise view is taken of meaning postulates, they sell the logicians' pass.

This is well illustrated by following through the implications of the processor design adumbrated in a recent SRI report. (As I am one of the authors of this report, I should make it clear that I am using this design as a vehicle for discussion, and not in a particularist critical spirit.) The report proposal adopts the logicist approach outlined earlier, for the purpose of building language processing interfaces to, for example, advisor systems. The design is for two processors. The first, the linguistic processor proper, is a general-purpose, application-independent component for syntactic analysis and the correlated construction of logical forms. The output of the linguistic processor is then fully interpreted (progressing from a representation of a sentence to that of an utterance) in relation to a discourse and domain context. The semantic operations of the linguistic processor proper deal, respectively, with the logical correlates of linguistic terms and expressions, and with the application of selection restrictions. The logical structures for linguistic expressions are determined by the domain-independent properties of items like articles and modals and of syntactic constructs like verb phrases, and by the formal characterisation of domain lexical items primarily as predicates of so many arguments. The lexical information about sorts, which supports the selection restrictions, is functionally distinct, as its role is simply to eliminate interpretations.

In this scheme of things the semantic information given for lexical items, and especially for 'content words', in the processor's output sentence representation, is fairly minimal. It is sparse, abstract, and opaque. The assumption is that predicates corresponding to substantive words are primarily given meaning by the domain description, and hence by the world which models this. Within the linguistic processor one sense of "supply", call it 'supply1', just means SUPPLY, where SUPPLY is an undefined predicate label. The sortal information bearing on predicate arguments which is exploited via selection restrictions does not appear in the linguistic processor's output meaning representation.

The domain description gives meaning to the predicates through the link provided by meaning postulates. These establish relations between predicates of the domain description language. But they are in a material sense part of the domain description, since these names are also used in the description of the properties of the domain world. Broadly speaking, the meaning postulates form part of the axiomatic apparatus of the domain description. Thus from a conventional point of view the lexicon says rather little about meaning: it merely points into a store of information about the world about which the system reasons, both to understand

what is being said and to react to this both in task appropriate actions and more specifically in linguistic response. The system structure is thus a particular manifestation of AI's emphasis on world knowledge and inference on this. The fact that meaning postulates are also the source of the sortal information applied through selection restrictions underlines the somewhat ambiguous character that meaning postulates have; but as noted, this sortal information does not figure as part of the information supplied in the representation of input text items in the output of the strictly linguistic processor. However the predicate labels of the meaning postulates may be word sense names so, e.g. 'supply1' is directly mapped onto 'provide3': this suggests that the boundary between semantic information in some narrow linguistic sense which refers to the content of the lexicon that is transmitted by the first processor, and semantic or conceptual information in the broader sense of the knowledge about the world that is incorporated in the non-linguistic domain description, has no theoretical but only operational status.

The immediate motivation for the system design just outlined is a very practical one, that of maximising system portability. Given our current inability to handle more than a very small universe of discourse computationally, we have to allow for the fact that some of the particular domain information appropriate to one specific application may be unhelpful or even confusing for another, and that the system design should therefore clearly separate the body of information which is general to language use and so should be transportable from that which is not. In the scheme presented the domain dependent information is confined to the lexical entries for the application vocabulary, and to the domain description.

But logicians also appear to advocate this form of processor as a matter of principle. Setting aside the question of whether the control structure of the processor is psychologically plausible (because it would be perfectly possible to apply syntactic and semantic, and linguistic and non-linguistic, operations concurrently), there is still the question of whether a viable general-purpose computational language processing system can be built with a strategy that treats meaning in the way the design described does, with so little information about it in the lexicon and so much in the knowledge base. The strategy implies both that there are no particular processing problems which would stem from the need to include both common and specialised knowledge, and perhaps several areas of specialised knowledge, in the knowledge base and, more importantly that none follow from the comparative lack of semantic information of the conventional kind found in ordinary dictionary definitions in the lexicon used for the purely linguistic processes, i.e. that there is only sortal information for selection restriction purposes. The first problem is not unique to the logicist position: any attempt to use information about the world, as all systems must, has to tackle the problem of arbitrarily related subworlds. The second problem seems to be more narrowly one for the logicist. The point here is not so much that, in staged processing, the attempt to avoid duplicating information means that information is unhelpfully withheld from earlier processes in favour of later ones. The point is rather whether, even in a situation where concurrent processing is done, providing much of the information germane to word meanings via domain descriptions is the right way to do semantics. It may be a mistake to regard linguistic meaning and world reference as the same; it is possible that some information about meaning has to be supplied, for representational use, in a form exclusively designed for strictly linguistic processing.

But all this is speculation. What is clear, on the other hand, is that the meaning postulates strategy, even if it does not involve the problems just mentioned, will, when applied to a non-trivial universe of discourse, imply vast amounts of very miscellaneous stuff. If language is intrinsically complex, simplicity at one point simply pushes all the complexity (or mess) somewhere else. In more exclusively linguistic approaches to processing this tends to take the form of putting all the detail in the lexicon: then the grammar can be nice and straightforward. Maybe one can get away with a simple syntactic analyser and semantic interpreter; but only by supplying all the specialised matter they need to work effectively on the various texts they will encounter, through lexical entries. And if simplicity in one place is found at the expense of complexity in another, it does not obviously follow that the system as a whole has the elegance of its simpler part. In the same way, in the logicist approach, the set of meaning postulates required may turn out to be such a huge heterogenous mass as to suggest, to the disinterested observer, that the purity the use of logic might imply has been compromised. From this point of view, indeed, whether the kind of information captured by meaning postulates is deemed to be part of the domain description, or is deemed part of the lexicon and is even expressed in the representation delivered by the linguistic processor, is irrelevant. Either way, the logicist approach suffers from the Same UnManageable Problem of miscellaneous linguistically-relevant detail as every other approach to language processing.

This is without considering the proposition that there is a much larger problem for which the logicists have so far offered us no real solutions: how to capture language use as this is a matter of salience, plausibility, metaphor, and the like. But whether or not the logicists can solve this, the real problem, we should not assume that they have got more mundane, literal matters of language taped.

H. Alshawi, R.C. Moore, S.G. Pulman and K. Sparck Jones, Feasibility study for a research programme in natural-language processing, Final Report Project ECC-1437, SRI International, Cambridge, August 1986.

J.R. Hobbs and R.C. Moore (eds), Formal theories of the commonsense world, Norwood, NJ: Ablex, 1985.