What's in a symbol: ontology and the surface of language - a Dialogue

Sergei Nirenburg Yorick Wilks

INTRODUCTION

YW: Researchers in natural language processing (NLP) all deal in non-numerical symbols: some treat texts as mere strings, most as objects in representations. The central issues for us are: first, whether or not one believes the symbols in representations are fundamentally language-like in nature, and, secondly, whether or not the answer to this question affects our expectations concerning development of large-scale representational resources for NLP (lexicons, knowledge-bases etc.) largely by automatic means.

If there were no relationship between our enquiry here into the nature of symbols and the processes within which we intend to use them, then our enterprise would be purely philosophical. The second issue above is currently of great practical importance in NLP. But we will argue here that the former, more apparently philosophical, question may influence the outcome of the research program.

While generally holding similar positions on most underlying issues, we are not in full agreement on the answers to these questions and what we say therefore has a dialectical tension or, if you prefer, has the implicit internal structure of a dialogue. We hope to make the points of disagreement clear as we go.

SN: Of course, even our agreement might be more of the kind illustrated by the following Soviet-era joke of the "Radio Yerevan" series:

A: Have you heard that Academician Ambartsumian has just won a Volga car in state lottery?

B: OF COURSE I HAVE. Only he's no academician. He's a night watchman. And his name is not Ambartsumian. It's Rabinovich. And it was not a car. It was a hundred rubles. And he played skat, not state lottery. Oh, and by the way: he did not win!

Some form of representationalism remains the mainstream AI position, even after much pressure from connectionism and, in the particular case of NLP, the recent and much publicized successes of purely statistical methods. For example, the purely statistical machine translation (MT) has risen to a level of success, in terms of the percentage of sentences correctly translated----of roughly the level of code redundancy in natural languages, i.e. 50-60%.

Our examination of the basis of meaning representation will return constantly to the following thesis: ARE REPRESENTATIONS IN NLP, KR ETC. CODED IN TERMS CLOSE TO THOSE OF NATURAL LANGUAGE ITSELF, and what are the consequences and perils of this fact, if it is one.

Of course, the crucial issue is to define what is meant by "close", or, in other words, what is natural language and where the boundaries between natural and artificial languages are drawn.

YW: We shall ask the question therefore, what are the salient features of the methodology that underwrite this ability of NLP and KR researchers to go on writing down knowledge structures, linguistic or otherwise, in formalisms whose abstract structure is defined, more or less, but whose content --- predicates, primitives, classifiers etc., --- are never set out in any formal terms but the most trivial?

SN: Here one has to agree first on the semantics of "formal"? Are there gradations of formality? Or do we all have to subscribe to logic and algebraic definitions of theories in order to be considered formal? Maybe the ability to be directly computable is what makes a structure formal?

YW: In any case, this question can only be asked of those who remain, more or less, within the representational paradigm: it cannot be asked of a fully distributed connectionist, nor of one, like Ken Church or Peter Brown who adopts a view of NLP in which words are no more nor less than symbol strings without significance.

In our day jobs we subscribe to the methodology of "NLP-as-Language-Engineering", a position that does not differ from connectionist and statistical NLP researchers. Where there is a clear distinction is with formalists who consider themselves, like Chomsky, to be doing science, not engineering. This is a complex and much discussed matter, and our position can be expressed very succinctly: formalists do indeed change formalisms but they never, ever, go through a process of rejecting a complex set of hypotheses in the face of large scale, statistically assessable evidence. That process, the distinctively scientific process, the lifeblood of science in all its aspects (except perhaps cosmology) never occurs in pure formal linguistics.

SN: Our paper, as will become clear, is in a very conservative tradition: core AI. It retains Newell and Simon's assumptions about semantics as information processing as much as it does McCarthy's vision of AI as the method of heuristics, as opposed to continuing attempts to make properties of natural language structures provable in some strong and unrealistic sense. Indeed some version of this dialogue gets published as a paper every five years or so, normally by an NLP AI-er, in which the mystery of natural language symbols as themselves representational for AI is turned over again and examined (Woods, Hobbs, Hirst have all done this recently), as well as AI commentators who find it not fascinating but faintly scandalous (Charniak and McDermott).

Within the representationalist camp, we wish to separate ourselves from those aspects of the formalist movement, whether within linguistics or mainstream AI, who believe the solution to whatever problem there is here, is to continue to seek formalisms with a logical semantics.

YW: We have discussed elsewhere the claims of the formal approach to NLP and will not repeat them here: in a nutshell, there is no reason to believe that systems for which notions like deductive closure are important have any demonstrable relationship to NLP, either as an empirical, engineering task or as a model of human processing.

Our discussion, necessarily sketchy, will be organized round the following five topics, or questions:

1. Are representation languages (RLs) natural languages (NLs) in any respect?

2. Are languages (natural or representational) necessarily extensible, as to sense?

3. Are language acquisition and extensibility linked?

4. If acquisition is possible, what are the consequences for any RL/NL difference?

5. The consequences of all of this for representations for humans (versus for machines).

SN: We cheerfully admit, of course, that we are unlikely to solve any of those problems in this discussion, but at least we hope to tease out additional issues and maybe to learn to formulate the difficult questions a bit better.

THE FIRST QUESTION: ARE REPRESENTATION LANGUAGE NATURAL LANGUAGES?

YW: This dialogue, though not a philosophical or psychological one, has clear overlap with one of the main aspects of Fodor's Language of Thought claims: that the basis of mental representation is language-like in nature. Fodor sets out a set of claims concerning the language-like properties of his putative LOT: in particular, the hierarchical, or tree-like, nature of its structures and the non-compositionality of the meanings of its predicates. The former has involved Fodor in extended disputes with connectionists about whether or not tree-structures can be replicated by connectionist learning techniques. The latter runs in the face of standard compositionality- oriented accounts of text (or sentence) meaning. We differ from Fodor on the crucial issue of what it means to be "language-like".

SN: Fodor's is only one set of criteria for what features make a language "NL-like". We can suggest additional ones. Two basic features, will probably include the "live", "unconstructed" character of NL and the functional criterion of being designed to support human communication, that is, relying on the human apparatus of understanding and thus gaining brevity.

YW: The first feature of language that should concern us in this discussion is as follows: can predicates of a representational language avoid ending up ambiguous as to sense? The negative answer to this question would make RLs NL-like. It will also mean that understanding a representation involves knowing what sense a symbol is being used in. If NLs are necessarily extensible as to sense ---and words get new senses all the time--then can RLs that use NL symbols avoid this fate?

SN: The predicates of a representational language are consciously CONSTRUCTED. They don't exist other than because of the will of the acquirer. We can argue about the process of construction and how the elements of a representational language get realized in practice. But the crucial difference is that NLs HAPPEN, RLs are MADE. You presuppose somehow that the RL is not constructed but rather EXISTS. And if, indeed, RL symbols are allowed to be ambiguous, then having to know in what sense the RL symbol simply sends the task of disambiguation one step further -- either to yet another, this time, unambiguous, RL or to one of the disambiguation schemata which do not rely on representing meaning in machine-tractable form.

YW: Here's just where we start to disagree strongly: for me, RLs are not made, or rather they are made up of existing NL bits--all too often English. And I can give no sense to the claim that we make symbols ambiguous or not--we have no such control of NLs or RLs.

SN: In some NLP applications, texts in an RL (such as the Mikrokosmos TMR) are typically used to represent the meaning of an NL text. Stating that RL elements are ambiguous is equivalent to saying that NL meaning cannot be truly extracted and represented. Another complication is the difficulty in using ambiguous RL statements as inputs to generation.

One can reconstruct the impetus for the above question in a deep pessimism about whether one can in fact construct an unambiguous RL. This is an important issue, though different from the original one. Briefly, there are two ways in which a representation language can be ambiguous -- first, when one and the same RL representation can correspond to two or more non-synonymous NL texts and, secondly, when one and the same NL text can be represented with two or more distinct representations. In the latter case, an added issue is how to establish whether these distinct representations are synonymous or differ only in the grain size of description (which might be considered allowable variation for the purposes of NLP applications). I believe that occurrences of ambiguity of the former are to be avoided if possible in RLs.

YW: Yes, but that's not what I mean by ambiguity in RLs. I see no difference between what you describe and the question of whether a passage of, say, Italian and of French are synonymous. I have no problem allowing that they are, modulo Quinean doubts. My worry is about the symbols that comprise them, be they RLs or NLs. Fodor faces this problem no more than do formalists who write runs(John) and appear to simply know which of the may senses of "run" the symbol bears in that context.

SN: One reason is that formalists do not usually work with an RL which has interpreted vocabulary (they sometimes refer to such an entity --- when they talk about models in model-theoretic semantics). Of course, runs(john) is not an expression in a natural language even if "run" is taken as an unspecified sense of the English word. The reason is almost trivial: the parentheses and the intended assignment of a function-like character to runs and argument-like character to john. At the same time, the artificial language used in this notation assumes an interpreter with human semantic capacity because a particular sense of the English word run must be selected, as well as a particular sense of john. Of course, the most blatant abuse of the similarity with NL and the most overt proof that such notations expect human interpreters is the ending -s on the function. Logicians would not think twice of writing run(John&Mary) fully expecting the interpreter to understand that runs and run are identical! It is only in this sense we can say that this representation is like English. The formal system in which such a language is used never bothers formally explaining this reliance on the human processor concentrating instead on studying formal manipulations of symbols.

YW: Your run/runs point is a nice one and tells on my side, I feel, at least as to how formalists and would-be formalists actually use formulae in a casual, self-deceptive, way. Shall we now ask, what then are the essential properties of being language-like and does a representational language have any of those properties accidentally or necessarily?

SN: NL RL ambiguous as to sense SN:+ YW:+ SN:- YW:+ extensible SN:+ YW:+ SN:+/- YW:+ constructed (vs. accepted) SN:- YW:- SN:+ YW:+ structures

hierarchical (JF) non-compositional presupposes human processor: -- at processing time SN:+ YW:+ SN:- YW:- -- at acquisition time SN:+ YW:+ SN:+ YW:- (?) primitives are: NL words/phrases SN:+ YW:+ SN:- YW:+ NL word-senses SN:- YW:? SN:+/- YW:-

YW: Let us put this matter very crudely: you seem to believe that the classificatory hierarchy of, say, WordNet consists of English words while that of Mikrokosmos does not. To me they seem both to print out an ascending thesaurus of similar terms up to an agreed Omega or top point, but I can see no difference between them IN PRINCIPLE only in richness of structure. This is simply a point about the interpretation of symbols in a verbal/ontology hierarchy and how they are normally interpreted.

SN: Normally interpreted by WHOM? By NLP researchers? Or by computer programs?

One of the differences between us may simply be the idea about HOW INFORMATION IS STORED in the static knowledge sources of an NLP system. I want to distinguish the following as separate static knowledge sources: the meaning representation (the actual RL text, not the formalism; consisting largely of meaning primitive instances organized into a structure), the ontology (ideally, a language-neutral store of meaning primitives) and the lexicon, one for each language, connecting NL elements with RL elements. The purpose of the meaning representation is to store, if possible, unambiguous statements which could be translated into a natural language (NL generation) and which could have been obtained by analyzing the meaning of an NL text or, say, by constructing a discourse as a reply to a database query. The system must carry out inferences on representations.

This capability would be easier to achieve if the symbols in the representation were unambiguous. If the representation is ambiguous, as I believe it would if RL were an NL, then the inference system would have to disambiguate RL symbols. The decision to retain ambiguity in RL leads to the situation where disambiguation occurs AFTER representation is obtained.

YW: I can accept of course that life would be simpler, if duller, if NLs consisted of unambiguous symbols, and the same for RLs, what I cannot grasp here is that it is, as you say, a matter of decision, to let a representation be ambiguous or not--I cannot understand that. Let me ask it at its simplest: how can you think the elements in Mikrokosmos, or Schank's CD, or anywhere else like that, are other than English words, with their own sense ambiguity?

SN: This question may be understood in at least two different ways. Let first comment on the one for which I feel I have a better repartee: surely RL elements are not just additional senses of NL words with which they share the ASCII codes. You wouldn't say that the Spanish MAYOR is another sense of the English MAYOR, would you? As to the second interpretation of your statement, let me just say that in a good RL every symbol is interpreted by its properties (relations to other concepts and/or values of measured properties). Unless grain size of description is considered tantamount to word sense difference, a putative different RL sense will be immediately detected, as its constraints will clash with those of the "old" sense.

If there is a separate language for primitives then it is easier to explain paradigmatic relations among word- and phrase-senses, such as synonymy, antonymy etc. Of course, as in WordNet one can bypass this explicitness about relations through the use of devices like synsets, but

then one ends up with a knowledge source which does not support all the operations necessary for automatic meaning processing.

YW: OK, so we are firmly in Quinean territory now, as when he used the impossibility of veridical paraphrase and translation to attack the folk-content notion of meaning. It would seem natural if we want to defend the latter (what do we call it--Commonsense semantics?) that we also believe in some provable/demonstrable notion of paraphrase. I would argue that even bad IE and MT are steps towards a practical notion of paraphrase--but that's not a philosophical defense.

SN: Do you mean that in order to defend the feasibility of meaning representation one needs to defend the feasibility of paraphrasing and translation? Clearly, even paraphrases and translations made by humans do not always convey EXACTLY the same meaning. The simple defense might be that philosophers habitually operate with an ideal RL and do not take into account the notion of the grain size of description, to say nothing about the possibility of an outright error or at least a slip of judgment in devising the content of an RL.

There is, however, a venerable tradition of describing meaning through translation and paraphrase without representing it. It is Frege who seems to have wanted a functional notion of word meaning representation (Sinn) that related surface entities (coreferents, for Frege) but which did not ITSELF YIELD/POSSESS CONTENT. For Frege an intension does not contain a coded meaning--it is just a function that allows you to specify or locate plausible referents in the world---a black box, a sort of recognizer (as opposed to parser) if you like!

And that brings us back to the issue of whether a representation of text meaning is required or can simply be pointed at and compared with the meaning of another text OR being connected with a particular denotatum (which is in the real world and not another set of symbols). In reality, we MODEL the world of denotata with a set of symbols because we cannot do otherwise if we want to put together many kinds of computer applications which require meaning for high quality.

YW: Well, of course, that is exactly what connectionists deny---they think they can give some sense to non-symbolic model--but I don't suppose we need bother with them here. You are right to bring in Frege: I suppose one could say that my obsessive questions about the exact status of primitives in a KR (and whether they are NL words or not) ignores his best known injunction not to consider the sense of the symbol, outside an expression. I naturally agree with the spirit of your defense of what is really the good-old-AI representationalist position, in defense of Frege, as it were, but can I indulge my skepticism for a moment more so as to wrap up this section. Can we get any closer to resolution here by considering a possible scale of NL likeness for an RL:

English or Bulgarian Esperanto Predicate calculus Some Interlingua

are these all equally expressive and if so how could we know or prove it? If they are then that is one NL/RL link--anything one can say in one one can say in the other. Certainly many users of Predicate calculus and Interlingual formalisms have held this position on equal expressivity.

SN: An RL must support automatic inferencing operations. One might just consider the difficulty (or otherwise) of adapting any of the above kinds of RLs to this task.

The major consideration is, again, whether the language is intended for people or for machines. The answer is easy in the case of English, Bulgarian and Esperanto. It is more problematic for predicate calculus. And it is impossible to comment on the last choice: interlinguae can be constructed in a very different fashion, just like any two independently constructed grammars of a language will be different (unless we'd have another case of the Septuagint) though may well have the same weak or even strong generative capacity. The ideal interlingua would be good to both computers and people: it would support inferencing in a broad domain, thus permitting high-quality meaning representation for texts; and it would also be easy to repair and expand (which, for the foreseeable future will remain largely a task for people).

YW: I am not sure this tells one way or the other on the language likeness issue, though it does make one ask if we have a good notion of "equivalent coverage" for representational languages in the way we do for grammars. So let me ask it this way: if our representation IS NL or anything like it does that interfere with a possible axiomatisation of any theory using it? How much does this require us to show our structures have reproducibility?

SN: How is reproducibility defined here? That another acquirer will acquire the same ontology and lexicon entries? This is patently unattainable, I am afraid. Is anything really at stake for us on the axiomatization issue? This would be, in part, an argument for the "science-like" character of our work--but as you have said axiomatization does not render a field scientific in and by itself. It's falsifiability that does.

YW: A key phrase that may help clarify our difference is that you say NLs are comprised of words and RLs of word-senses. We may want to make the point that language study is DIFFERENT from other AI areas: because IN ALL CASES BUT LANGUAGE, we can imagine a computer system being better than us--better than physicians or grandmaster chess players. We CANNOT imagine the system understanding better than people, and this point is not often appreciated in some NLP areas.

SN: Well, what about general perception or motor skills? But back to our topic. I may want to say that the purpose of a representation is to get the symbol ambiguity OUT--which is EXACTLY what you think cannot be done. But does that point require an objective measure of symbol ambiguity--and IS THERE ONE--anywhere in our discussion or outside it?

YW: Somewhere in his discussion of what he calls The Concrete Lexicon versus The Abstract Dictionary Martin Kay seemed to be arguing that the brain MUST subscript symbols to separate the senses of the RL primitives, and for him that would be in the Concrete Dictionary, i.e. the head. I have never been sure quite what he meant, but he was clearly discussing the same issue as us, as many have before, and he seemed to me to be on roughly my side here: conceding that the RL atoms could be ambiguous and something would have to be done about it by the processor that used them--the brain itself. Notice that this is not the same as the RL expressions being ambiguous which is again what you take me to mean here and I do not.

SN: If indeed we are talking about the brain, I am agnostic. I don't know an awful lot about what is going on inside that device.

YW: No, neither do I, but people like us who talk about the nature of RLs for human knowledge must, like all AI workers, be making potential claims about the brain, whether they admit it or not.

SN: I wonder whether we indeed do. Maybe if we concentrate on representation by computers and for computers, we will be off the hook.

The second question: are languages necessarily extensible, via ambiguity propagation and resolution?

YW: How can anything that is a language be other than extensible? If that is obvious, one can then ask how can such extended information about it be acquired. This can be seen as a traditional Chomskyan question about language and the child's learning of L1, but we intend it in the more accessible sense of an enquiry about how a computer can come to acquire information about language, and could that ever be equated with the mastery of a merely finite, static, resource.

SN: Of course, language is extensible. However, any sublanguage used in an application, has, up till now, been finite and static. In AI applications, acquisition of knowledge typically precedes its use. And when a new word must be entered in the lexicon of an MT system, it has been done by people or, at least, sanctioned by people. One can argue that the associated representation language was again static and was used on any new text as such, until the need for further extensions arose. In other words, there exists a decoupling between acquisition and use, while in NLs when use as it were extends itself.

YW: Another of our key questions here is whether this feature of language is UNIVERSAL and, if so, MUST IT BE ALSO BE POSSESSED BY REPRESENTATIONAL LANGUAGES. We shall then move from answers to those questions to the central question of whether we also want to say that THE SYMBOLS OF A REPRESENTATIONAL LANGUAGE FUNCTION LIKE THOSE OF A NATURAL LANGUAGE in this particular regard.

SN: One little consideration here against strong similarity between RL and NL symbols as regards the property of ambiguity. It is well known that people have difficulty recognizing ambiguities; they immediately choose the contextually appropriate sense for each word or phrase. This seems to suggest that, if indeed meanings are represented, the elements of the representation are not ambiguous, as the operation of retrieving the other senses of an input language element is so expensive.

YW: Ah, yes, this is Wittgenstein's good old "the senses of a word do not pass in front of my mind" point. But this does not, to me, prove anything about the nature of the representation -- it is a point about our lack of ACCESS to our processes. And in any case I am not claiming that representations are ambiguous --only that the items in them CAN BE ambiguous (out of context presumably) in just the way NL items can--so this point does not touch me at all. A difference of emphasis between us reflects our intellectual upbringings, I fear, you focus on the whole representation (in RL), I on the RLs constituents!

Do we therefore need to discuss the issue of what it is to know, or assess, objectively, in some sense, that a symbol in a representational system/language is AMBIGUOUS. It is clear from

the variation of lexicographic intuitions (10 senses for a word vs 2, in different dictionaries) that mere intuition is not enough. Consider, too, Wierzbicka's argument that polysemy is mostly an illusion.

SN: Surely, lexicographic intuitions are about NL, not necessarily RL. That lexicographers disagree may simply mean that there does not exist some "correct" number of senses. I intuitively dislike the suggestion, but maybe in some system-operational approach, one could define word senses cross-linguistically. This latter point connects with the idea of using the target language as RL for a source language in MT in those systems that don't go for machine-oriented meaning representation. It also connects with the idea of using this almost Hjelmslevian view of the semes across languages as an impetus for humans to select senses for representation even in an internalized RL, such as, e.g., the Mikrokosmos text meaning representation.

YW: Yes, the translation as representation case, between NLs, has had a new lease of life recently hasn't it, and it is a strange shadow of the Fodorian comedy of the LOT as the translation one can't get at. I used to suffer a temptation at meetings to ask him how he KNEW the LOT WASN'T Italian, or something like that, but I fortunately never gave way to it--since I know he doesn't know.

More seriously, given that LISP was considered almost a LOT by AIers in the seventies---consider NIL in LISP, now usually thought of as 3-ways ambiguous (an empty list, an atom, and a Boolean value). Was there an objective test of that claim? Did it matter until it was noticed--in terms of the usefulness or otherwise of LISP? Was there a formal criterion for spotting it--i.e. is "giving a formal semantics of a representation" a revelatory mechanism for exposing "ambiguity"? I suspect not, and so much the worse for giving formal accounts of things!

SN: The fact that a lexical ambiguity that "happened" in a representation language ended up being contextually "benign" does not necessarily prove that ambiguity can be introduced with similar impunity into RLs designed for the purpose of representing meanings of texts. Another interesting fact is that the empty list, (), an expression (phrase) in LISP, is always interchangeable with the atom NIL, which is a word. In object-oriented LISP NIL is an object of type NULL. Is, then, NULL an element of an RL serving another language (in this case, LISP, not an NL?).

YW: No, types aren't IN formal languages--whereas in NLs the types (e.g. animate) can of course be in the NL because as we know, NLs can be metalanguages for themselves--this is probably a point on your side showing that NLs and RLs differ on this point. Though I still do not need to concede, as you insist, that we allow or prevent ambiguity in RLs. These matters are under no one's control, not in LISP, and certainly not in RLs like CYC where no one could control the coders' use of the predicates. There is no RL/NL distinction there, where you seem to want it for RL coding, and this, for me, rebuts your earlier claim that applications are static and finite.

The case of corpus statistics may be interesting here because its users generally have no use for terms like "word sense" which they find unbearably intuitive; for them, symbols simply occur in environments which may or may not be usefully separable into classes of occurrence.

I am not sure there is any objective demonstration of the ambiguity of a symbol--i.e. the Reality of Word Senses? I have always used the Schvaneveldt Pathfinder nets as a justification; they can show "bank" having separable subgraphs with an algorithm that requires no seeding or stimulation to do that. The other well-known statistical methods usually do not show ambiguity unless you assume it to start with.

SN: First of all, as an aside, we need to remind ourselves that the lexical-semantic issue of word and phrase senses and their disambiguation is only a component of the general problem of RL. Composition of senses and representing nonpropositional meaning are among the important additional components of the general issue. Concretely, boundaries of word senses, as usually recorded in dictionaries, are notoriously fluid. No two dictionaries agree on the exact number and content of senses of a lexeme. There are different attitudes to sense determination. Not only Wierzbicka but also the "lexical-rule"-oriented lexical semanticists prefer to propose few (usually, one) word sense for recording in the dictionary and then to add rules for accommodating meanings that do not directly conform to statically defined constraints. I see this as an exercise which is nice but not essential, either in theoretical or practical sense.

YW: Here is a slightly different approach: the relation between an expression in NL and its corresponding RL may be a relationship like that between a language and its metalanguage or one of (presumably mutual) translation. If the former case holds, then the languages do not really differ in type--they simply have an asymmetric relationship and might differ in expressiveness, but, as is well known, the ML/RL is as much in need of ITS ML as the NL/object language. There is an agreement in the formal world to stop worrying about this, and probably rightly, but, if the relationship is of that sort, there is no reason to believe the two levels differ over, say, polysemousness or extensibility of term meanings.

Alternatively, if the relationship is translation, then, almost by definition, TRANSLATE (X, Y) if X and Y are both symbolic, and not, say, traffic lights, requires that X and Y be of the same TYPE, i.e. NLlike, in this case!

SN: We cannot get into an infinite regression of metalanguages. The relation between NL and RL is, of course, asymmetrical, though there will be both many to one relations between elements of NL and RL (e.g., synonymy) and one to many ones (most notably, polysemy). Internal consistency is achieved for RL through maintaining the complex cross-relationships in the ontology (the RL vocabulary). The issue of meaning grounding is more difficult and we might want to state, cautiously, that it is achieved in the Fregean sense, through the multiple connections of elements of an RL with multiple NLs, through human judgement of quality of translation (correspondence). Incidentally, it will be easier to write NL understanders and generators if the RL does not allow paraphrases.

Your argument about the relation of translation hinges centrally on how one defines TYPE. It might be that we don't disagree but you elect to stress similarities between NL and RL while I persist in looking for differences. Let them be of the same type, but RLs must support machine inferencing while this cannot be asked of NLs. The case of the Dutch company BSO working on Esperanto as interlingua clearly shows how much a human-oriented (though invented) language had to be modified in order to serve as a kind of RL. Even the developers themselves, Esperanto enthusiasts all, had to call the new language somewhat differently: BCE or "binary-coded Esperanto."

==========

THE THIRD QUESTION: ACQUISITION AND EXTENSIBILITY ARE THEY LINKED?

YW: Acquisition in our sense is linked to the necessity or otherwise of symbol ambiguity, because much acquisition is of new ambiguities or senses of symbols.

SN: The extent to which automatic acquisition of content is plausible may indeed be a major practical undercurrent of this paper. A question for you: does explaining the meaning of an ambiguous symbol in terms of another ambiguous symbol actually constitute disambiguation?

YW: This a practical question, too, of course. We are seeking, in our normal day jobs, and outside dialogues like this, practical, robust, NL processors, not necessarily wedded to one particular theory, but ones that tackle areas of NL and KR representations. I am, in a sense, rather neutral about representation-- -modulo the very general representation issue at the core of this discussion--but strong on assessment and large systems and data. On your question: again, I accept that an (ambiguous) symbol can be defined, more or less, by a definitions string that is not, as a whole, ambiguous.

An assumption about communication behind all this is that the trivial diagram we are all familiar with of humans communicating with their separate representations via the very narrow linear language stream from their mouths, is wrong in one crucial respect. It is normally shown with the SAME structure in the two heads. But there is no reason at all to believe that human communication requires identical logics, lexicons, grammars, parsers etc. in both heads, any more than it does identical beliefs. I suggest the most striking feature of communication is that humans who differ about these structures can communicate, just as can individuals with different languages, dialects, etc. multi-lingualism and problems of differing beliefs and lexicons WITHIN A SINGLE LANGUAGE are no different at all on this view; nor is communication diachronically between earlier and later users of the same language.

SN: Yes, there should be no presupposition of a similarity between the knowledge and processing resources of various people, modulo of the hardware (wetware?) and possibly some other (genetic?) constraints. The difference is clear in the case of conversations between people who are native speakers of different languages (cf. the point below about the Japanese using English as interlingua), belonging to different professional and social strata, people of different ages, etc. It is indeed amazing how adaptable people are when viewed as information processors. At the same time, on the surface, what this shows is only that there may be as many "proprietary" devices for processing language as there are people.

YW: The commonsense fact is that communication can take place within a bandwidth of difference, and human-computer communication in a way explores the limits of this bandwidth and how far it can be extended in special cases by tuning lexicon structures and beliefs to each other in the course of communication itself. But this issue cannot be separated from the problem of language representation itself-- e.g. we cannot understand the nature of the representation of meaning in lexicons, say, unless we can see how to extend lexicons in the presence of incoming data that does not fit the lexicon we started with. Extension of representation is part of an adequate theory of representation.

SN: I think I understand your intended meaning: first, no set of static knowledge sources will have complete coverage; therefore, representations need to be extensible; therefore there must be a mechanism of adding elements to representations, preferably, on the fly.

Further, many of such representation elements are lexical. And the easiest way of naming these new elements would be through the natural language strings that refer to them in the input (which triggered the representation augmentation process in the first place).

This, of course, presupposes automatic acquisition, because if a human is involved in acquisition other suggestions could become quite palatable. In short, the argument for natural language in the representation becomes thus also practical: we need it because otherwise we'll have problems naming new atoms (and we would not want the gensym solution for reasons discussed below within the Fifth question).

YW: Suppose we write

I: structure1 X corpus --> structure2

as a basic model of acquisition of a representational structure, be it an ontology or a lexicon, to indicate that a state of the structure itself plays a role in the acquisition, of which structure2 is then a proper extension (capturing new concepts, senses etc). This is a different model from the wholly automatic model of lexicon acquisition in, say TIPSTER related work, which can be written:

II: corpus --> structure

the latter case is one which does not update or "tune" an existing lexicon but derives one directly and automatically from a corpus. We are arguing the essential role of representational structure in this process, and hence the first process -- which we may also take to involve some essential human intervention as well. But whatever is the case about that we are not discussing the ab initio / tabula rasa case. Interestingly perhaps, neither of these is an NLP analogue to the Chomskyan approach to language acquisition, which might be written:

III: constraints X corpus --> structure

Since the constraints are of the same type as a structure this form is closer to I above, especially in Fodor's work, where the constraints become a form of ur-ontology or ur-lexicon.

SN: This classification seems to skirt the issue of human involvement. In reality, fully automatic acquisition of lexical information does not, at this time, go anywhere deep enough to allow use in hard problems such as full-text lexical disambiguation or even syntactic analysis. In TIPSTER, for instance, as far as I know, the automatic acquisition of subcategorization patterns for some English verbs was accompanied by massive manual acquisition. Personally, I would choose to use a combination of all three of the above methods of acquisition, depending on the quality of the input data and availability of good-quality constraints and structures. In general, I feel that automatic acquisition is an open research problem equivalent in complexity to general NLP, and current results in automated acquisition can be considered achievements only when the rules of the game are modified (e.g., when the task is defined as disambiguating a few carefully selected words relative to the set of senses in a particular printed dictionary).

==========

THE FOURTH QUESTION: IF ACQUISITION IS POSSIBLE, WHAT ARE THE CONSEQUENCES FOR OR AGAINST RLs AS NLs? [OR ---IS NL ITS OWN METALANGUAGE?]

YW: If automatic acquisition of content is possible to any degree, from an MRD or corpus then, since those are plainly in NL, does this suggest that in some form NL is a representation language for information about language, and that settles the metalanguage possibility raised earlier.

SN: First of all, I think that the premise is more or less a moot point at this time, because automatic acquisition of content can be deemed possible only if content is defined in plainly trivial terms. Any success in the automatic acquisition of content is predicated on the ability of the developers to model (in the weak sense, with no claims of similarity of the model to the modeled other than at output!) the disambiguation and other meaning assignment processes of humans. More concretely, this modeling involves overt, human-directed, formulation, at time of acquisition, of the background knowledge and processes which support the automatic assignment of meaning at processing time.

But even if the premise of your argument is given, the argument as a whole still seems to me a fallacy or at least a sleight of hand. It is rather similar methodologically to the use by our colleagues at USC ISI of the fact that the ontology in the Pangloss MT project used English as its atom metalanguage (in that the Spanish lexicon explained the meanings of Spanish words in terms of an ontology whose atoms were homographs of English words and expressions) to turn the Spanish-to-ontology lexicon into a Spanish-English bilingual dictionary used in lexical selection for generation.

YW: Well, if they can do it, I might want to say it is not a sleight of hand but proof of my NL-RL point. I also want to use the metaphor of a dictionary as containing a lexicographer's "conscious", which is what we might extract by these processes--but other computations over the result could yield meaning connections no lexicographer had actually seen (and which might be said to model his unconscious). How do you relate this to full explicitness for MIKROKOSMOS coders ---you assume must they have thought of everything?

SN: The answer to the last question is purely methodological: the acquirer does not have to produce all the knowledge deductively. A version of the code-test-debug loop is applied instead. In general, I have a much more prosaic view of the acquisition. The content of an MRD is in NL and this *IS* a representation but this representation is good for people and bad for machines.

It needs to be reformulated and FLESHED OUT for machines -- as lexicographers in writing dictionary entries heavily (if subconsciously!) rely on the fact that their representations, i.e., definitions will be processed by a high-quality language processor, namely, the human! This may be the crux of our disagreement... The task of NLP knowledge acquirers is to use their language processing capacity to state information as overtly as possible given a desired grain-size of description AND in a format which facilitates access by machine (e.g., frames). The latter condition is, of course, of secondary importance: it is a convenience consideration only.

The former condition is contentful in that it presumes that the definition is not complete by itself but only together with the human understander of that definition. This can be proven wrong, incidentally, if it is shown that dictionary entries, in fact, do not rely on extraneous human knowledge in specifying definitions. But if that were so, why do lexicogrpahers say that if you don't know some meaning, you won't understand it from the dictionary? Is this just frivolity?

YW: I still think it is an open question whether structures derived pretty much automatically from MRDs can be useful for NLP. If they are your position weakens. Our difference is really one of bottom up vs top down approaches to the same information? You in fact believes that the ACQUISITION of the core of these knowledge sources can be done semi-automatically, but UNDER HUMAN SUPERVISION, for instance, in automatic production of lexicon entries through LEXICAL RULES.

THE FIFTH QUESTION: THE RELATIONSHIP OF THESE ISSUES TO REPRESENTATIONS FOR HUMANS AND FOR MACHINES.

SN: I believe that, in fact, there is a MAJOR difference in writing for people and writing for machines. Of course, the purpose of the description of a meaning should play a role (e.g., whether this is for general knowledge or to teach a machine or a human to assemble a pump). However, the representations for humans (see above) assume the presence of an extremely powerful analysis system and a huge amount of background knowledge. One has to specify things at a much finer grain of description for machines than for humans even if the purposes of the two descriptions are compatible. YW: This is an excellent question and not as much discussed as it should be. A difference in machine vs. human handling of representations is over what one might call the GENSYM issue---a machine can handle English expressed as arbitrary Gensyms for words, a native speaker cannot without vast retraining, if then. We can both accept the difference -- between the comprehensible representations that humans need and the fact that they have no meaning for machines--and use it to prove opposite views as regards NL and RL. Your observation proves to me that, for that very reason, RLs must be accessible to humans (as well as machines) and THEREFORE must be NL like in certain respects.

SN: This is a weaker form of your original argument about NLs as RLs, and I would fully agree with the premise: just like the computer programs, which are written in part to be read by people (an estimated 80% of the time of software engineers is spent on maintenance --- that is, reading and improving other peoples' programs), so should the knowledge structures in an RL. That in, say, Mikrokosmos names of atoms are words or phrases in English is due exactly to this desideratum. It is, on my view, a conceptual fallacy to read more into this state of affairs---for instance, to claim that there is an intrinsic necessity for RL elements to be also elements of an NL. Of course, sometimes there can be side effects, as in the uses of the SENSUS "ontology" in Pangloss which was mentioned earlier.

Charniak's argument against connectionism was that you could not UNDERSTAND the structures they acquired; and they were therefore not acceptable representations, regardless of whether machines could use them. How much of our underlying disagreement is over whether structure must be comprehensible?

SN: Well, to comprehend anything which is non-trivial, one must learn. One can, in fact, learn to read meaning representations. It has been proved in practice. Of course, it is very desirable to avoid having people read unadorned RL structures, but this might be a premature hope.
YW: Would we be helped by thinking about how actual coders use RLs? The example that has interested me most is that of how many Japanese researchers use interlinguas for, say, MT or in the EDR dictionary project, but with English symbols. It has been argued that it may be a positive advantage for them because they do not, in many cases, see more than the main sense for any primitive and this makes its use easier and less confusing than for a "native speaker" of the interlingua, if you will allow that term. The question might then be: has that any analogies with how you see a machine as handling a representation--the difference between human and machine handling of representations being, I think, crucial for your position, though not for mine?

SN: I know how you feel about being handicapped in international fora (where English is, volens nolens, a lingua franca) by virtue of being a native speaker of English. Without a doubt, this rather clips one's rhetorical wings because, motivated by the desire of being understood, one would select what one believes would be language understandable by one's audience. Of course, not being a native speaker, I do not feel the loss as acutely. Or maybe I believe that obscurity breeds respect...

The analogy with understanding by machines is also clear --- after all, they operate with fewer word and phrase senses (to say nothing about connotations) than people. However, I do not see any bearing whatsoever that this observation has on the differences between RLs and NLs. If the abovementioned Japanese researchers do not know English well enough, this does not impinge in any way on the issue that RLs for computers should be either bad or good English or any other NL, or an artificial language (with either narrow or broad coverage of meanings in an NL).

YW: Maybe when we model understanding we aim at too high a target: in ordinary situations people may understand just a fraction of what is said by a speaker but they ask clarification questions only when IT MATTERS. In reality, there are few penalties for failure, i.e. miscommunication or misunderstanding: contrast medical counseling dialogue, legal searches, patents, and philosophical discussions, in all of which misunderstanding is thought disastrous and maybe carries real costs.

SN: This is one of the issues on which we have a complete agreement. In practical terms, however, the mechanisms for understanding what is important and what is less so in understanding text are probably as complex as any problem in NLP and AI. It is probably easier to tell the machine to understand everything it can without ranking the elements of input in importance.

YW: I have a feeling we may have swapped sides here a bit. If I understand the "empty-representation / concept /intension / TMR" point above, you seem to accept that "symbol-free linking" role as being the ideal for representations but one unobtainable at present so we must make do with the forms of weak self-representation that this discussion is about. I think I am more Wittgensteinian and think language central and unreplaceable in thought and representations so that there will never be any alternative to doing what we do now --whatever happens to AI/CL---because we are self-defined by language and we can't expunge it from representations.

SN: If the issue here is that, however people might try, they won't be able to produce RLs which are not ambiguous, it is verifiable. Possibly, this will be happening asymptotically. But it is surely not plausible that people are somehow constitutionally UNABLE to come up with an unambiguous RL, that is, not due to the size and complexity of the problem (which can be ameliorated through tools, partial automation etc.) but rather by definition! We need to go through a much more detailed discussion of the influence of the fallibility of human acquirers on the nature of the RL -- Sapir-Whorfian influences of native tongues, difficulties with listing all senses of a lexeme or all synonyms of a word, as opposed to great facility of judgment whether any two words are synonymous (or stand in any other static semantic relation.

YW: Ah, so we do differ--I think it is beyond human ability to design an RL without the features they now have, and for the reason we touched on: they must remain comprehensible to us, and if they do they will be like NL ---where like is still a bit fuzzy. It is as if comprehensibility will carry a price--the price of loss of control of the sort you think we can retain.

SN: Ah, but it is exactly your understanding of the meaning of "like" which is the crux of the matter here! As long as it is fuzzy, one cannot very well argue about it. Further, I assume that by "control" you mean whether people can be taught deliberately to produce representations that, as they or their project managers believe, would be processable by machines. As I understand it, you think this implausible because vestiges of human language will prevail and corrupt the RL representations. I think it necessary, unless we can teach machines to reason using knowledge bases which are inconsistent or ambiguous. Mind you, I do not have any illusions about the practical attainability of knowledge bases which are fully consistent and non-ambiguous. The methodological choice is to carry on pretending that they are until special mechanisms are developed for dealing with such inconsistencies and ambiguities.

YW: Correction--vestiges of NL in an RL does not mean for me that a machine cannot "understand" it--of course not, it is happening all the time in a thousand working programs.

SN: Right. As long as we agree on what constitutes understanding. But that can be the topic of another conversation!