WHAT CAN PHILOSOPHY CONTRIBUTE TO THE STUDY OF NATURAL LANGUAGE PROCESSING?

Martin Ringle
Computer Science
Department
Vassar College
Poughkeepsie, NY
12601

For the past twenty years philosophers have observed the development of research in natural language processing (NLP) and have offered periodic critiques of both its methods and its goals. (See Bar-Hillel, 1964; Matson, 1976; Dreyfus, 1978; Searle, 1980; and Odell, 1981.) Much of the criticism has proven to be valuable and artificial intelligence workers such as Winograd (1980) and Woods (1981) have acknowledged the positive influence of philosophical input to their work.

A great deal of philosophical criticism of natural language processing (and of artificial intelligence in general) however, rests, on misconceptions about the actual goals and claims of this research. This is partly due to the fact that NLP workers have not explicitly established a set of methods and aims for their work; it is also due to the fact that there are actually a number of different goals which motivate NLP research.

The purpose of this paper is to spell out the different objectives in the field of natural language processing in order to identify the places where philosophical criticism is legitimate and useful as well as those areas where it is inappropriate. Hopefully, this analysis will be valuable to philosophers and AI workers alike.

Research in natural language processing can easily be misconstrued to be a concerted effort towards a single goal. In the simplest terms, this goal would be the implementation of a system whose linguistic powers matched those of a literate, native user of a natural language such as English or French. In fact, however, research in natural language processing is a loose amalgam of projects aimed at a variety of goals. Even though common research requirements exist, such as the development of techniques for parsing, inference, memory organization, and so forth, presuppositions, methodologies, and criteria of success differ in significant ways from one project to the next. It is somewhat misleading, therefore, to appraise or to criticize the theoretical foundations of natural language processing as a single enterprise. Yet some philosophers (e.g., Odell, 1981) have assumed that the principal goal of NLP is the unified goal just mentioned, and have proceeded to question the plausibility of the research on that ground alone.

Consider the following formulations of the aims of natural language processing research:

1. To design systems which will allow a user to perform some traditional operation(s) on a computer (such as database query) without thereby requiring the user to learn an artificial language or a set of formal constraints which must be applied to the use of natural language.

2. To design systems which will be capable of processing textual material in order to produce accurate summaries, reliable translations, (from one natural language to another) or stylistically acceptable prose.

3. To design systems which will permit the user to initiate and direct a dialogue, in natural language, in a particular topic domain, with the latitude and fluency available in ordinary human dialogue.

4. To design systems which will persuasively exhibit the full range of human linguistic abilities, such as reading, translating, paraphrasing, interrogating, conversing, and so on.

5. To design systems which are able to use and understand natural language in precisely the same way that people do.

6. To design systems whose workings provide us with an explanatory model of the structures and processes responsible for human language use and understanding.

The first four formulations involve pragmatic goals, the fifth represents an epistemological goal, and the sixth an explanatory goal. The vast majority of efforts in natural language processing fall into the category of pragmatic goals. (See Waltz, 1982 for descriptive surveys of recent NLP projects.) Most, in fact, are examples of the first or Type 1, goal. Systems such as LIFER (Hendrix, 1977), ROBOT (Harris, 1979) and LUNAR (Woods, Kaplan, and Nash-Weber, 1972), for instance, provide natural language front-ends which are used principally for database query. The research aims which motivate the construction of these systems (and others like them) are relatively modest insofar as the use of natural language is constrained by topic, vocabulary, syntactic breadth and user dialogue goals.

The Type 2 goal is slightly more ambitious, since the analysis, generation, or translation of text may require a system to deal with a broad range of topics, a large vocabulary, complex syntactic constructions, and the intentions of an author (or reader) which may be less than obvious. Progress towards this goal has not been as substantial as progress towards the first goal, but there are programs which can analyze and paraphrase text (DeJong, 1982), produce modest translations from one natural language to another (Wilks, 1973) and generate moderately smooth English prose from an internal semantic representation (Mann & Moore, 1981).

Serious efforts towards the Type 3 goal are very few in number and have appeared only within the past five years.

Systems in this category include SRI's TDUS (Robinson, 1980) and BBN's HWIM (Bruce, 1982). Neither these, nor other systems of this sort, have achieved a level of combined reliability and efficiency which would make them suitable for broad implementation. However, there has recently been

a great deal of attention turned towards this area and a greater effort to achieve this goal can be expected in the near future.

The Type 4 goal is one which has been popularized in science fiction and the lay press, but it is not cited by AI researchers as the rationale for any serious NLP programming effort. This is not to say, of course, that AI workers have not entertained the idea of such a goal as a backdrop for their activities. In the proper perspective, such a goal is analogous to the one which underlies physics (and the natural sciences) namely, the eventual discovery of all lawful relationships among natural objects. Physics, after all, is dedicated to the objective of ultimately explaining the universe in terms of quantitative laws. One does not, however, invoke this goal as the aim of any particular research project. Moreover, it would be absurd to try to criticize a particular line of research in physics by attempting to show that this long-range goal is untenable. Even if the universe is not ultimately knowable in terms of the principles of physics, the enterprise still provides us with an ever-increasing understanding of natural phenomena. The same holds true for research in natural language processing: Even if the long-range goal is unattainable --- and that remains to be shown---this does not affect the plausibility of the other three pragmatic goals nor does it invalidate the knowledge of natural language processing derived from programs designed to achieve those goals.

The fifth goal raises a completely different set of questions. Here we are concerned with the status of the performance rather than with the performance itself. In the case of pragmatic goals, the criterion of success is the degree to which a system is able to deal effectively with linguistic input (or output). The phrase "deal effectively with" may be interpreted differently for different applications, but in general it implies that the system is able to carry out a function which would involve use and understanding of natural language if performed by a human. The claim is not made, however, that the system actually uses or understands natural language itself. We can appreciate the point of this last statement by considering the following question: Can the pragmatic goals be pursued without pursuing the epistemological goal as well?

Some AI researchers would undoubtedly say 'yes' in answer to this question and would point to the success of natural language interfaces such as the one used in MYCIN (Shortliffe, 1976), which are not generally characterized as "language understanding" systems. Cautious researchers, such as Winograd (1973) and Leitner (1977) have emphasized the epistemological limitations of their programs by putting the word "understand" in quotation marks when using it to refer to their natural language systems.

Other researchers, however, freely speak of their programs as natural language understanders. Schank and Riesbeck, for example, go one step further and argue that natural language programs must be directed towards genuine understanding:

> Computer programs that attempt to replicate understanding without simulating the human understanding process are doomed to failure when it comes to very complex processes. Nowhere has this been clearer than in natural language processing (Schank & Riesbeck, 1981, p. 2).

The point that Schank and Riesbeck make is a crucial one. If we are concerned with a Type 1 pragmatic goal, then genuine understanding is probably superfluous. A Type 1 interface can be limited to such a well-defined area of natural language that we can design systems to "deal effectively with" the range of anticipated linguistic input by means of deterministic production rules, discrimination nets, or similar methods. But if we are interested in Type 2, 3, or 4 pragmatic goals, then we must accept the fact that the potential for novelty, diversity, and deviant usage of linguistic inputs may be so great that a system would be effective under such conditions only if it were able to process the meanings of those inputs. And this implies that it must be able to genuinely understand natural language.

It follows, then, that while the Type 5 goal may be irrelevant to the majority of pragmatic systems of the present (and recent past), it is essentially related to the development of the more ambitious pragmatic systems of the future. It is in this context that philosophical evaluations of natural language processing become relevant: by analyzing the conceptual requirements of genuine language understanding, the philosopher can illuminate the theoretical conditions which an NLP system must meet. Moreover, unless these conditions are met, the epistemological goal cannot be achieved and thus the more ambitious pragmatic goals cannot be realized. Whether or not AI workers explicitly view the Type 5 goal as a motivating force in their research, therefore, they must acknowledge its indirect relevance if they intend to pursue a Type 2, 3, or 4 pragmatic goal.

The Type 6 goal is one which has drawn a great deal of attention in artificial intelligence due to statements such as the following:

> We consider the theory and model of semantic nets to be a computational theory of superficial verbal understanding in humans (Simmons, 1973, p. 63).
>
> . . .[W]e shall describe a model of human language understanding that forms the basis for a set of computer programs. . .(Schank, 1973, p. 187).

Both of these statements were published nearly ten years ago and since then there has been a considerable change in the claims made for the psychological significance of AI programs. Nevertheless, some AI researchers (especially members of the Yale Group) still view the explanatory goal (Type 6) as a primary one, and some philosophers (e.g., T. Simon, 1979) still find the view to be worthy of criticism.

An argument to demonstrate the relevance of the Type 6 goal to the rest of natural language processing might go something like this:

> Genuine understanding is necessary for any natural language understanding system capable of achieving Type 2, 3, or 4 pragmatic goals.
> Genuine understanding can be achieved only by processing language in the same way that humans process language.
> A system which does things in the same way as humans do them can serve as a model for explaining human language processing.
> Therefore: Pursuit of goal Types 2 - 5 entails pursuit of goal Type 6.

There are, however, several problems with such an argument. The second premise asserts a "process-product" identity relation which is very

much open to dispute. There are numerous instances (e.g., the synthesis of urea) where an artificial process results in a substance, event, or function which is identical to a natural substance, event, or function in every respect save its mode of origin. It has yet to be shown that a cognitive ability, such as the understanding of natural language, can be produced only by employing exactly the same processes and structures which are involved in human language understanding. Indeed, it has yet to be conclusively shown that all human beings understand language by means of exactly the same processes and structures.

However, even if we accept the second premise ---under some interpretation of the phrase "in the same way"---the conclusion still does not follow. A program for natural language processing is not, itself, an explanation of anything. In order to be viewed as explanatory, the details of a program---its variables and data structures, its control structures, and so on---must be interpreted with respect to human processes and structures. Any program can be legitimately interpreted in a variety of ways, few of which will bear any relation to the concerns of human psychology. The explanatory value of a natural language program, therefore, is not inherent in the program itself but arises, rather, from the use which can be made of it by someone who is concerned with cognitive modeling. The use of AI programs to theorize about human language processes, in fact, is not AI research at all. It is a tool of cognitive psychology or, if one prefers, a methodological heuristic for a multidisciplinary investigation of phenomena such as discourse comprehension, text comprehension, and so forth. It does not follow, therefore, that research in natural language processing entails explanatory goals of Type 6; consequently, philosophical objections to AI programs as theories of human language abilities are irrelevant to the plausibility of AI research in natural language.

Of all the types of goals ascribed to natural language research, then, philosophical evaluation is directly pertinent only to the epistemological goal formulated as Type 5, above. More specifically, the only valid judgment philosophy can provide is one which says that "the concept of natural language understanding entails X, hence a system must (be, do or have) X or it will not be capable of natural language understanding." The only valid objection philosophy can make to natural language processing is that "a computer, in principle, cannot (be, do or have) X."

REFERENCES

Bar-Hillel, Y. "The Present State of Automatic Translation of Language." In F. L. Alt (ed) Advances in Computers, New York, Academic Press, 1964.

Bruce, B. C. "Natural Communication Between Person and Computer." In Lehnert & Ringle, pp. 55-88.

De Jong, G. "An Overview of the FRUMP System." In Lehnert & Ringle, pp. 149-176.

Dreyfus, H. WHAT COMPUTERS CAN'T DO, Revised Edition, New York, Harper & Row, 1978.

Harris, L. "Experience with ROBOT in Twelve Commercial Natural Language Database Query Applications," IJCAI Proceedings, 1979, 6: 365-368.

Hendrix, G. "The LIFER Manual," SRI Technical Note No. 138, 1977.

Lehnert, W. & Ringle, M. (eds) STRATEGIES FOR NATURAL LANGUAGE PROCESSING, Hillsdale, NJ, LEA, Inc., 1982.

Leitner, H. "The Determination and Conceptual Structuring of Restricted Domains of Discourse for 'Intelligent' Interactive Systems," SIGART Newsletter, 1977, 61: 51-52.

Mann, W. C. & Moore, J. A. "Computer Generation of Multiparagraph English Text," American Journal of Computational Linguistics, 1981, 7: 17-29.

Matson, W. SENTIENCE, Berkeley, California, University of California Press, 1976.

Odell, S. J. "Are Natural Language Interfaces Possible?" *IBM Systems Research Technical Report* TR73-024, 1981.

Robinson, A. "Understanding Natural Language Utterances in Dialogs About Tasks," SRI Technical Note No. 210, 1980.

Schank, R. "Identification of Conceptualizations Underlying Natural Language." In Schank & Colby, pp. 187-247.

Schank, R. & Colby, K. (eds) COMPUTERS MODELS OF THOUGHT AND LANGUAGE, San Francisco, W.H. Freeman, 1973.

Schank, R. & Riesbeck, C., INSIDE COMPUTER UNDERSTANDING, Hillsdale, NJ, LEA, Inc., 1981.

Searle, J. "Minds, Brains, and Programs," The Behavioral and Brain Sciences, 1980, 3: 417-457.

Shortliffe, E. H. COMPUTER-BASED MEDICAL CONSULTATIONS: MYCIN, New York, North Holland, 1976.

Simmons, R. "Semantic Networks: Their Computation and Use for Understanding English Sentences." In Schank & Colby, pp. 63-113.

Simon, T. W. "Philosophical Objections to Programs as Theories." In M. Ringle (ed) PHILOSOPHICAL PERSPECTIVES IN ARTIFICIAL INTELLIGENCE, New Jersey, Humanities Press, 1979.

Waltz, D. "The State of the Art in Natural Language Processing." In Lehnert & Ringle, pp. 3-32.

Wilks, Y. "An Artificial Approach to Machine Translation." In Schank & Colby, pp. 114-151.

Winograd, T. "A Procedural Model of Language Understanding." In Schank & Colby, pp. 152-186.

Winograd, T. "What Does it Mean to Understand Language?," Cognitive Science, 1980, 4: 209-241.

Woods, W., Kaplan, R. & Nash-Weber, B. "The Lunar Sciences Natural Language Information System: Final Report," BBN Report No. 2378, 1972.