

PROSPECTS FOR PRACTICAL NATURAL LANGUAGE SYSTEMS

Larry R. Harris
Artificial Intelligence Corporation
Newton Centre, Mass. 02159

As the author of a "practical" NL data base query system, one of the suggested topics for this panel is of particular interest to me. The issue of what hurdles remain before NL systems become practical strikes particularly close to home. As someone with a more pragmatic view of NL processing, my feeling is, not surprisingly, that we already have the capability to construct practical NL systems. Significant enhancement of existing man-machine communication is possible within the current NL technology if we set our sights appropriately and are willing to take the additional effort to craft systems actually worthy of being used. The missing link isn't a utopian parsing algorithm yet to be discovered. The hurdles to practical NL systems are of a much more conventional variety that require, as Edison said, more perspiration than inspiration.

It should be clear that none of my remarks conflict with the obvious fact that NL research has miles to go and that there are innumerable unresolved issues that will continue to require research beyond the foreseeable future. Our understanding of NL has merely scratched the surface, and it is fair to say that we don't even understand what all the problems are, muchless their solution. But by using the powerful techniques that have already resulted from NL research in extremely restricted micro-worlds it is possible to attain a high enough level of performance to be of practical value to a significant user community. It is these highly specialized systems that can be made practical using the existing technology.

I will not speculate on when a general NL capability will become practical, nor will I speculate on whether the creation of practical specialized systems will contribute to the creation of a more general capability. The fact that there is a clear need for improved man-machine communication and that current specialized systems can be built to meet that need, is reason enough to construct them.

The issue of whether practical specialized NL systems can now be built is, in my opinion, not a debatable issue. Those of us on this panel and other researchers in the field, simply don't have the right to determine whether a system is practical. Only the users of such a system can make that determination. Only a user can decide whether the NL capability constitutes sufficient added value to be deemed practical. Only a user can decide if the system's frequency of inappropriate response is sufficiently low to be deemed practical. Only a user can decide whether the overall NL interaction, taken in toto, offers enough benefits over alternative formal interactions to be deemed practical.

If we accept my point that practicality is in the eyes of the user, then we are led to the inescapable conclusion that practical NL systems can now be built, because several commercial users of such a system [Pruitt, O'Donnell] have gone on record stating that the

NL capability within the confines of data base query is of significant practical value in their environment. These statements plus the fact that a substantial body of users employ NL data base query in daily productive use clearly meets the spirit of a "practical" NL system.

The main point of my remarks is not to debate the semantics of practicality, but to point out that whatever level of utility has been achieved, is due only in small part to the sophistication of the NL component. The utility comes primarily from a custom fitting of the NL component to the exact requirements of the domain; and from the painstaking crafting of the lexicon and grammar to achieve the necessary density of linguistic coverage. In a sense, practicality is derived from a pragmatic approach that emphasizes proper performance on the vast bulk of rather uninteresting dialog, rather than focusing on the much smaller portion of intellectually challenging input. A NL system that is extremely robust within well-defined limitations is far more practical than a system of greater sophistication that has large gaps in the coverage.

Attaining this required level of robustness and density of linguistic coverage is not necessarily as intellectually challenging as basic research, nor is it necessarily even worthy of publication. But let's not kid ourselves -- it is absolutely necessary to achieve a practical capability! It has never been clear to me that members of the ACL were interested in practical NL systems, nor is it clear that they should be. But I think that it is fair to say that there aren't many practical NL systems because there aren't very many people trying to build them! I would estimate, on the basis of my experience, that it takes an absolute minimum of 2 years, and probably more like 3 years, to bring a successful research prototype NL system to the level of practicality. This "development" process is well known in virtually all scientific and engineering disciplines. It is only our naivete of software engineering that causes us to underestimate the magnitude of this process. I'm afraid the prospects for practical NL systems look bleak as long as we have many NL researchers and few NL developers.

Pruitt, J., "A user's experience with ROBOT," Proceedings of the Fourth Annual ADABAS User's Meeting, April, 1977.

O'Donnell, J., "Experience with ROBOT at DuPont," Natural Computer Conference Panel, May, 1980.

