2002: ANOTHER SCORE

David G. Hays
Metagram
25 Nagle Avenue, Apartment 3-G
New York, NY 10040

Twenty years is a long time to spend in prison, but it is a short time in intellectual history. In the few years just prior to the foundation of this Association, we had come from remarkably complex but nevertheless rather superficial analysis of text as strings of characters or, perhaps, lexical units to programs for parsing that operated on complex grammatical symbols but according to rather simple general principles; the programs could be independent of the language. And at the moment of foundation, we had--in the words of D. R. Swanson--run up against the stone wall of semantics. No one at the time could say whether it was the wall of a prison yard or another step in the old intellectual pyramid.

On my reading, the record is unmistakable. The best work of the past twenty years has been on a higher level than that of 1962. Those who learned about syntactic, semantic, and cognitive structures as students must feel quite scornful of the timidity with which we introduced these new topics to a world that doubted their propriety. But then some were not so timid. After all, the new ideas are in the curriculum.

Meanwhile, the commercial significance of strings of characters has come to everyone's attention. So-called word processors are widely used, and the market grows. Commercialization of our most rudimentary techniques has taken twenty years. We may wonder how long it will take to put on the market systems with our more recent, more advanced techniques, but we can be sure that the market will eventually buy them.

We can also be sure of encountering new barriers. Our most important gain in the past twenty years is, as I see it, the assurance that whatever barrier we meet can be climbed. This is no case of "Climb one, climb them all." Such arrogance is folly. Language is closely associated with thought. Knowledge of them both, and of their association, is just what carried us over the barriers that were insurmountable twenty years ago. The barriers we meet are inherent in the systems of thought that we use. We know enough about thought to announce that its characteristic and discriminating feature is the capacity to generate new and more powerful systems of its own kind. A railroad does not become an elevator when it reaches a cliff, but thought does just that.

No one anticipated in 1962 that the study of language or the investigation of "thinking machines" would lead in twenty years to an understanding of how intellectual barriers convert themselves into scaffolding for the erection of new theoretical systems, and no great social institution--not the university, and certainly not government--has yet recognized the revolutionary thrust of our small enterprise. The world understands, vaguely, that great change is taking place, but who understands that the pace of change will never slow down?

Intellectual progress consists in the routinization of the work of intuitive genius. Before the Renaissance in Europe, some persons by insight could establish the sum of two numbers or the truth of some fact about nature. Since the Renaissance we take these accomplishments so much for granted that we scarcely understand the system of thought in which they were problematic. At most twenty-five years ago, the determination of the surface structure of a sentence was problematic. By now we understand rather clearly how phonological, syntactic, semantic, and cognitive considerations interact in parsing.

We, as a global culture, have taken a step comparable to the Renaissance, and we, as the members of an Association, have had a significant role. Advances in linguistics, in cognitive science, in the art of computation, and in artificial intelligence have contributed to our work. Some would say that we are merely users of their results. I think that we have supplied a crucial element, and I understand our name--computational linguistics--to designate our special conceptualization.

Until we went to work, the standard conceptualization of analysis in western thought was translation into a universal scheme. Logic, mathematics, and computation assumed that all argument would be by manipulation of certain forms. The logician, mathematician, or computationist was expert in these forms and manipulations. Given any problem domain, someone would translate its material into the standard form. After manipulations, someone would translate the results back.

Computational linguistics has the idea that computation can be designed on the pattern of linguistic theory. In this respect, it seems to me, there is a sharp distinction between computational linguistics and natural language processing. The latter seems to belong to artificial intelligence, and artificial intelligence seems to be the inheritor of the standard assumptions. I think that computational linguistics has the advantage.

Language and thought are fantastically complex, and when their mechanisms are translated into the old universal forms the representations

are equally complex. Yet, from the right perspective, we have seen that language and thought have simple structures of their own. If we translate linguistic mechanisms into computational terms, the first step is hard, but the rest is comparatively easy.

The making of software is still, as it has been from the beginning, a grave problem. For this problem I see only one remedy. Computational mechanisms must be translated into the terms of the user for whom the rest will be easy. But the user is not unique; the class of users is heterogeneous. Hence computational mechanisms must be translated into many different kinds of terms, and so far this translation seems very difficult. "Metagramming" is my name for an approach to the simplification of the hard part.

For thousands or tens of thousands of years humanity has engaged in the translation of linguistic mechanisms into the terms of different perspectives on the world. Thus, cultures and languages vary in profound ways. And cultures and languages vary together. Until now no one has understood this process. It went on in billions of brains, and it was effective. Now we try to understand it and to extend it from the linguistic level to the computational.

The curious formula that we offer for the conversion of intellectual barriers into scaffolding is just this: Formulate a description of the barrier. Translate the mechanisms of thought or of computation into the terms of the description. Execute the new mechanisms. As I see the matter, such work was done by intuitive genius until recently, but we are routinizing it. This formula generalizes on a central notion of computational linguistics and seems to me our first contribution to universal knowledge.

The formula contains an inexplicit element. What are the terms of the description to be? In what language does one formulate the description? I see no plain answer to this question. In fact, I am willing to take it as identifying, but not as describing, the next barrier.

Another way to put the matter is to say that the proper description of the barrier is a metaphor of its elimination. Metaphor is at present in the same limelight that illuminated semantics twenty years ago. We have not yet found the correct angle to illuminate the problem of metaphor, the proper language for description of the problem.

Again, I suggest that metaphors serve us in discussions of abstract matters. Surmounting an intellectual barrier is stepping to a higher level of abstraction or, in a somewhat novel technical sense, moving to a metalevel.

And finally I point out our inability to characterize the mutual influence of any complex whole and its myriad parts. If we consider a play or novel, a religion, a culture, or a science and ask how the unique quality of the whole emerges from the mass of elements, we have little or nothing of a scientific nature to say. And if we ask how to construct a system of this kind, how to design a building or a programming language, how to enhance a culture or educate a child, we find ourselves with traditions and intuitions but without explicit theories.

So I see a goal worth scoring, and I imagine the possibility that computational linguistics can move toward it. Deep study of computation inculcates powerful methods of thought, and deep study of language supplies the right objects of thought. Computational linguistics contains what I reckon to be needed by those who would wrestle with abstraction, metaphor, and metasystems.

Mankind is a complex whole, and its individual human parts are myriad. The computer in every home will alter the mutual influence of person and population. For better or for worse, no one can yet say. Moral issues arise. Technical issues will determine not only whether morally sound principles can be put into practice, but also how we formulate the moral questions. Here is work for twenty years to come and beyond.