The Current Status of Computer Hardware and Software as it Affects the Development of High Quality Machine Translation

by

D. Walker

The Mitre Corporation

The developments in computer hardware and software over the past ten years have gone a substantial way toward satisfying the needs specified in the early '60's as prerequisites for effective machine translation programs. In particular, the storage capacities and processing speeds of current computers far exceed some of the stipulated requirements established during that period. Increases in sophistication of programming systems have paralleled hardware developments as evidenced in operating systems like OS for the IBM 360 and 370 series and Tenex for the PDF-10, to name only two. Compiler technology also has advanced markedly during the period, particularly as elaborations of the syntax-directed techniques introduced about ten years ago. Programming languages as well have increased in breadth, flexibility, and power, so that, although assembly language coding certainly still would reduce run-time, it no longer is a cost-effective alternative. As a result it seems reasonable to say that hardware and software considerations no longer constitute major obstacles to machine translation, at least according to strategies that are currently being pursued.

In spite of the conciseness—and, I believe, the essential correctness—of the foregoing summary statement, two observations need to be made before considering the implications I believe can be drawn from it. There are no systems for machine translation that I am aware of which use algorithms designed specifically to take advantage of recent computer capabilities. That is, the strategies currently pursued are those established in the early

1960's. While hardware and software may not be obstacles, it is not clear that they have been used to full advantage. However, looking at the other side of the issue, it also is not clear that new approaches, particularly those motivated by the recent concerns with semantic processing, might not result in specifications for machine architecture or programming that cannot be met by existing equipment and procedures.

Whatever importance is assigned to these observations, it is clear in any case that the problems of mechanical translation at this stage are primarily of two kinds, linguistic and algorithmic. That is, the responsibility for establishing hardware and software requirements depends on the design specifications for a mechanical translation system. And these specifications entail a knowledge of the grammars of the language involved, a strategy for analyzing them, and a procedure for relating the analyses. Until we can resolve these matters satisfactorily, any prescriptions for hardware and software are purely speculative.

In spite of these uncertainties, one class of computer capabilities should be stressed in this context both because of its potential use in the process of mechanical translation and because of the role it may play in grammar development and in the formulation of algorithms for linguistic analysis. I am referring to interactive capabilities that allow for on-line access to the computer. Although it is only recently that such man-machine systems have become cost-effective, it is still worth asking why machine-aided approaches to machine translation were not proposed and pursued in earlier years. Logically, they would seem easier to implement than

would fully automatic approaches. Again, I suspect that the problem here as before is the lack of understanding about grammar, linguistic analysis, and translation algorithms. However, there has been a substantial amount of work now with grammar testers and with systems for handling personal files, work that should be extended into the mechanical translation arena.