

Session 8: INFORMATION PROCESSING AND LINGUISTIC ANALYSIS

INTRODUCTION

EDMUNDSON: As you know, this Symposium is intended to concern itself primarily with machine translation. You will note that the sponsoring universities intended the Symposium primarily for professionals active in the field of machine translation, for linguists, and language scholars. We also felt, and therefore said rather clearly in the brochure, that the Symposium would be of interest to computer scientists, information retrieval specialists, librarians and documentalists, and manufacturers of data-processing equipment. We did this for two reasons. We wanted to make sure that all the research groups would come to the Symposium with the feeling that they could present research papers of a technical nature which would be discussed by their fellow workers active in research. On the other hand, we did not want to close the door to people working in other fields, such as information retrieval and library documentation. We welcome them. I hope that this session will point out some of the interrelations between this particular discipline called machine translation and some of the new areas that have also come about in this second phase of the computer revolution. The first phase was devoted primarily to numerical analysis. I think it is quite clear to all of us that the second phase is directed toward linguistic analysis. We have talked about the contributions that linguistics can make to the field of MT. We have also discussed some of the ideas that have appeared in MT research which may feed the area of linguistics. We have heard discussions of this in the previous panels. Now, what are the other areas? I am thinking particularly of what we might call automatic abstracting. What is the interrelation between automatic translation and automatic abstracting? Should we translate first, say from Russian to English, and then have automatic abstracting programs which abstract the English, or should we devise automatic abstracting routines in Russian and then translate the abstracts from Russian to English? Finally, how should we handle users' requests? There are some very interesting cost analyses that can be made between the former and the latter systems. I hope that the speakers will treat this in more detail. We are also concerned with the tools or the kinds of basic research that might be expected to

Session 8: INFORMATION PROCESSING AND LINGUISTIC ANALYSIS

help machine translation research. I am now thinking of the interdisciplinary nature of this field. We have, in this room, some of the best linguists, language scholars, mathematicians, statisticians, logicians, computer engineers, and information retrieval experts. However, it is not at all clear that we are going to get a great deal of help from mathematics and logics, even though it may be very convenient to make use of some of the notions of mathematics. It may very well turn out that concepts in algebra, topology, and statistics provide nothing more than a neutral ground to model linguistic phenomena. It may very well be that we will find ourselves, or even have found ourselves, in a position similar to that encountered in information theory. Perhaps it did not solve immediate problems but it did provide insight and a catalyst. Can we develop some catalysts for MT from the classical branches of mathematics, or do we have to create new disciplines? It may very well be that today's mathematics will not decide things for us. We do not have any constructive way of determining this. We do not even know if some of our linguistic problems are intrinsically undecidable, in the sense of logically undecidable propositions. With this in mind, I hope the speakers today; Dr. Thyllis Williams of the ITEK Corporation, Professor Oettinger of Harvard University, and Dr. Garvin of the Ramo-Wooldridge Laboratories; will exposit some of the results they have developed in the MT field. Also, I hope they will introduce some of the ideas they have regarding new basic research tools that will help us solve some of the problems in the general field of automatic analysis of linguistic data of which auto-translation is just one phase.