

[Presented at the MIT conference on mechanical translation, June 1952]

The Structure of the Problem of Mechanical Translation

Olaf Helmer
The RAND Corporation

The problem of mechanical translation has three principal components: (i) the formulation of a set of specifications for the objective to be attained; (ii) the design of a translating machine; and (iii) theoretical translation problems.

Decisions under (i) will have to be made under consideration of the purpose to which the translation output will be put. The prime distinction here is as to whether the emphasis is placed on an accurate transmittal of the cognitive content of the input or rather on a faithful rendering of the style and emotional content. In the former case one may want to require a strict logical equivalence of input and output, or one may be satisfied with a certain (high) degree of message equivalence, in which case it would be desirable to have a criterion by which it is possible to decide which of two proposed translation schemes is to be considered preferable.

As for (ii), it appears that highspeed general-purpose computing machines will be able to handle the main translation task. Engineering design therefore should concentrate on ancillary devices designed to reduce the input and output bottlenecks. In particular, an automatic reading machine would be necessary for fast mass translation. In addition, radical devices for reducing the input and output volume of the translation machine might be considered, such as automatic pre- or post-translation abstracting methods.

The theoretical translation problems, (iii), divide themselves into syntactical and semantic ones. The former are concerned with word and sentence structure, the latter with the meanings of words and phrases. Both probably require separate translation procedures. Syntactical problems, in principle, can be handled by a logical analysis of the languages in question, while semantic problems lead into questions of word frequency, optimal dictionary size, and the dependence of ambiguity of meaning on the context.