Book Reviews

An Introduction to Machine Translation

W. John Hutchins and Harold L. Somers
(University of East Anglia and University of Manchester Institute of
Science and Technology)

Reviewed by
Guðrún Magnúsdóttir
University of Gothenburg

An Introduction to Machine Translation covers the linguistic and computational technology of machine translation and gives a basic survey of relevant systems and research projects. It starts systematically with linguistics and the problematic aspects of understanding languages, and continues with accepted computational techniques for analyzing language. The chapter on basic strategies describes established architectures in system design. The chapter on analysis concentrates on the linguistic problems and the knowledge available to solve these. The chapter on transfer and interlingua describes contrastive problems concerning the translation of languages and how MT approaches them. Generation of natural language is briefly mentioned—understandably so, since MT has been an application where, with the exception of a few systems, generation has been reduced to the level of a morphological process. In the chapter "Practical use of MT systems," the modes of operation, or the way in which the user interacts with the system, are brought forth, not the actual use of a system in an end-user environment. The chapter on MT evaluation gives good insight into the many ways in which a system can be evaluated and the different values that have to be considered depending on who is evaluating what.

The last nine chapters of the book describe current systems and research projects. These chapters cover work on MT during the past as carried out in the Western world, regrettably neglecting Asian systems and research, even if that is in quantitative terms larger than all other work in MT combined. The last chapter briefly mentions new directions of research and gives perhaps the best guidelines for further reading. But here the suggestions are rather dated and the reader is well advised to look for more-recent publications by the authors who are mentioned in Section 18.10.

An introduction is seldom written without different opinions about what to include. The interested reader will always find that he or she wants more information, and achieving that must be the ultimate goal of an introduction. However, there is a contextual discrepancy in the book between the theoretical solutions put forth in the linguistic introduction and the practical solutions applied in the systems described. The authors maintain a focus on syntactic theory, whereas the practical systems that have been commercially successful are mostly nontheoretical in their syntactic approach and rely heavily on lexical pragmatics. Thus the book implies that syntactic theory is more established in MT than it really is. This is a controversial problem in MT research and development: on the one hand, how to make MT theoretically re-

spectable and on the other, how to prove that linguistics, mostly syntax, is the best technical approach for MT. This influences developers to be secretive about their system technology, and researchers to strive to make a theoretically 'correct' system, but without much success. In fact four of the eight systems described here in detail—SUSY, Ariane, Rosetta, and DLT—have not made the market in spite of excellent linguistic design. Assuming that MT implies some contrastive research, Eurotra has no bearing upon MT (Eurotra Assessment Panel Report 1993). Systran, in spite of its unorthodox linguistics, Météo, with its small domain, and Metal, with its extensive lexical coding (which is clearly understated in this book) are systems that are currently being used as practical systems. All share some sort of simplification as regards linguistic splendor.

Current statistical approaches to MT are very briefly mentioned. They deserve a more in-depth treatment, so that the reader may understand the simplicity of the approach and its limitations regarding improved translation quality beyond what is already achieved.

The shortcomings of this book, which are always subjective, do not outweigh its relevance. It is the best existing basic course book for an introduction to MT, for students who have a basic training in linguistics and, preferably, some knowledge of computer science. It is very well written and easy to read. The content of the book is systematic and well structured. The references to further reading at the end of each chapter are extremely useful. For lecturers giving basic courses on machine translation, this is a long-needed addition to the literature.

Reference   Report 1993 (in press).
CEC (Commission of the European
   Communities) Eurotra Assessment Panel

Guðrún Magnúsdóttir is a director of studies in natural language processing at the University of Gothenburg, and has given several courses on machine translation and worked on surveys and evaluation of MT. Her address is: Institutionen för spraakvetenskaplig databehandling, Göteborgs Universitet, S-412 98 Göteborg, Sweden; e-mail: gudrun@svenska.gu.se.