

ances that carry "false information." But more importantly for the claim about the distinctiveness of HPSG, we do not see substantial interactions between semantic and syntactic phenomena. This contrasts with GB theory, for example, in which the subtleties of quantifier interaction are supposed to depend in a very direct way on details of syntactic structure, and at least some of the semantic rules operate under the same substantial constraints as syntactic rules. In HPSG, on the other hand, the "flow" of semantic information is defined by a special semantic principle tailored to fit the needs of the fragment considered here. The work done so far could just as well have been done as an afterthought. However, if the promissory notes are redeemed in Volume 2, I expect that the semantics will play a larger role.

NOTES

1. They mention the "speculative" solution that simply disjoins the increasing obliqueness constraint with a constraint saying that $-N$ constituents precede focused constituents, but this idea obviously needs further development to work even on the range of cases considered in the text. Pollard and Sag refer to technical reports by Uszkoreit on this problem.
2. The contrast is not clearly formulated. For example, Pollard and Sag note that whereas the first-order formulas *laugh(rebecca)* and *run(rebecca)* may both denote the same truth value (in the actual world at a time), the formulas $\langle\langle\text{laugh, laugher:rebecca}; 1\rangle\rangle$ and $\langle\langle\text{run, runner:rebecca}; 1\rangle\rangle$ (in the actual world at a time) will always be "more contentful." In an introduction, though, it is worth considering the clear sense in which the first-order formulas, like the sentences *Rebecca laughs* and *Rebecca runs*, have more content: they assert (under the intended interpretation) something about the world, whereas the others (under the intended interpretation) simply denote abstract objects without telling us *anything* true or false. What is the motivation for going to the lengths of saying that a situation in which the circumstance holds is a fact? Furthermore, the latter expressions denote different circumstances only if the *run* relation is different from the *laugh* relation, and it would be useful, even in an introduction, to alert a student to the reasons that defining appropriate identity conditions on these relations is a very tricky business. The situation is not clarified by Pollard and Sag's further suggestion that while $\langle\langle\text{believe, believer:claire, believed:}\langle\langle\text{laugh, laugher:rebecca}; 1\rangle\rangle; 1\rangle\rangle$ is well formed, the first-order formula *believe-(claire, laugh (rebecca))* is syntactically ill-formed. This is not even correct, since *laugh* can be both a predicate and a function in a first-order language. In fact, we can define the function *laugh* in such a way that *laugh(rebecca)* denotes the very fact of a situation in which the circumstance denoted by $\langle\langle\text{laugh, laugher:rebecca}; 1\rangle\rangle$ holds. The real issues are missed without a slightly more careful development.

Edward P. Stabler, Jr. is an assistant professor of computer science at the University of Western Ontario and an Institute Scholar in the Artificial Intelligence and Robotics Program of the Canadian Institute for Advanced Research. He received his Ph.D. in philosophy from MIT, and has recently been working on parsing as deduction and transparent implementations of GB theories. Stabler's address is: Department of Computer Science, University of Western Ontario, London, Ontario N6A 5B7, Canada. E-mail: stabler@uwo.uunet

MACHINE TRANSLATION SYSTEMS

Jonathan Slocum (ed.)

Cambridge, England: Cambridge University Press, 1988, ix + 341 pp.

(Studies in natural language processing)
ISBN 0-521-35166-9, \$49.50 (hb); ISBN 0-521-35963-5, \$16.95 (sb) [20% discount to ACL members]

Reviewed by
John S. White
Planning Research Corporation

Machine translation systems is a successful attempt at presenting the breadth of issues on machine translation from the most relevant of perspectives, namely the systems that exist. The orientation toward presenting the research platform systems, the prototypical systems, and the systems in use enhances the current efforts toward finding the common ground between researchers and implementers. Such convergence leads to fresh insight for the implementers, and, for the researchers, solutions to the practical but vexing problems that the production systems have already solved.

The papers in this volume are a new presentation of articles in the special two-issue *Computational Linguistics* coverage of machine translation. There have been some updates to the content of these articles, though more updates would have painted a more accurate picture about changes, for better and worse, in the fortunes of these systems.

Though there is no explicit explanation of the format of the articles, it is apparent that they were written in accordance with some suggested outline or questionnaire. Thus the heading numbering and organization of the articles are roughly parallel. The advantage of this organization is, of course, that the different systems can be readily compared on the basis of design, theory, and performance. The disadvantage is that there is a tendency to respond to the guidelines without giving a clear indication of what the guidelines were.

The papers in the volume are the following:

Jonathan Slocum, "A survey of machine translation: its history, current status, and future prospects"

This paper is a version of the invited paper Slocum presented at the 1984 COLING conference at Stanford. It is a valuable statement about machine translation, and one which could bear up well with periodic updated republication. The theme is that an understanding of the issues of machine translation presupposes an understanding of translation itself. The need for translation, the way in which professional human translation is done today, and therefore the way that machine translation approaches fit in, should be critical components in any machine translation design. Yet it frequently is not,

which is why so many such systems fail not just from poor performance but from poor user acceptance.

The descriptive axes Slocum uses have become the standards for characterization of particular system designs and overall methodologies. He divides system function into MT (machine translation) and MAT (machine-assisted translation, divided in turn into machine-assisted human translation and human assisted machine-translation). On the other axis, the system design, he distinguishes the direct and indirect, the latter again into interlingua and transfer systems. These classifications are consistent with those made by others, like Garvin and Bruderer, yet expressed in a very accessible way. It is likely that it is this accessibility of writing style, as much as the clarity of the descriptive axes, that has made this a seminal paper in machine translation studies.

Slocum addresses both research and production systems in his discussions, which is atypical of survey articles (which tend to present either the research systems or the commercial systems but not both). There is a risk to such an undertaking in the commercial computer system development arena: information about status, goals, and even existence of such systems changes very quickly. The value for the purposes of comparative designs and approaches remains valid, however.

Axel Biewer, Christian Féneyrol, Johannes Ritzke, Erwin Stegentritt, "ASCOF: A modular multi-level system for French-German translation"

In ASCOF, Biewer et al present a research-oriented French-German system with the perspective that modular processes, corresponding to linguistic phenomena, be callable from any time the linguistic phenomenon occurs. Central to this discussion is their means of handling complex noun phrases, which is done as a process separate from the identification of simple phrases. There are some concerns here, for example, whether verbal coordination is done in an entirely different way from noun coordination (which loses generality for French analysis). Also, it is not clear how well simple verb phrases with complex noun phrase arguments are handled, as where a particular noun is not a suitable candidate for complementation of some verb, but the noun phrase of which it is head is valid. Nevertheless, the notion of re-use of the same process in different areas of analysis is worth discussion.

The ASCOF system is a relative of the University of Saarland SUSY system, and is a transfer-type design, employing a set of ATN subsystems for parsing. Transfer and synthesis of German is accomplished by a transformational grammar, though at the time of the original writing of the article the German synthesis had not been accomplished. An update on the success of the planned generation strategy would be helpful. Modularity of the lexical component is also an important part of

the ASCOF system, in which semantic relational information is not replicated for the lexical entries, but rather such information resides on a semantic network accessible to the lexicon.

Bernard Vauquois and Christian Boitet, "Automated translation at Grenoble University"

Among the more important aspects of the Grenoble machine translation work, as reported by the late Bernard Vauquois and Christian Boitet, have to do with maintenance of the optimum development environment in the context of a mainframe-based system. Many of the critical considerations of system design, from the component-specific, special-purpose languages, to the implementation programming language itself, revolve around the fact of relative inefficiency of available Lisp environments on the more traditional mainframe systems. While largely directed toward research and development, this aspect of the Grenoble systems has direct significance for commercial natural language systems, where the decision about whether to implement on the mainframe on which relevant data resides, or on a workstation networked to that mainframe, is a crucial one.

Begun in 1961, the early Grenoble system was an interlingua prototype for Russian-French translation. The current transfer-based system development began in 1972. A variety of smaller prototypes of other language pairs, as well as the GETA Russian-French system share a programming environment known as ARIANE-78, which also has user functions (pre- and post-editing). Some concern arises from the implication that ARIANE-78 has a special text file format, which is a problem for compatibility and transportability.

A system comparison with the Kyoto system, claimed to be similar in platform, compares such things as instruction size and storage requirements, without expressing the crucial comparisons, namely, speed and accuracy. The claim that an implementation in Lisp would be "40 times more voracious" again points to the issue above of opting to build and run a natural-language processing system directly on a mainframe rather than one of its clients: an inefficient Lisp running on the mainframe has consequences for possibly hundreds of users at runtime. The cost of avoiding this outcome is implementation in low-level languages, which makes maintenance and compatibility difficult.

The large variety of special purpose, component-specific development languages appear to be rather cumbersome for a linguist to use, and are presented in a way that makes for too many acronyms to try to track within the course of the paper.

Winfield Bennett and Jonathan Slocum, "The LRC machine translation system"

In the description of the University of Texas METAL system, Bennett and Slocum continue the theme of

orientation toward the whole translation task. METAL has been oriented toward production translation problems since 1979 (when Siemens first became directly involved in funding). As a consequence, developmental decision points that would be equal in a research environment have favored the directions that would make the process of translation efficient for the translator.

METAL is a transfer German–English system (there are also Dutch–French and German–Spanish pairs) employing an augmented phrase structure grammar whose analysis nodes can be decorated with node-specific transformations, transfer, and synthesis instructions pertinent to that node. The concept of the grammar rule as locus of analysis, anaphora, transfer, and generation activity has made for a development environment conducive both to linguists and to programmers. Inflectional morphology is also handled by the grammar, so that morpho-syntactic phenomena do not have to cross component boundaries. As is common with modular transfer systems, METAL has a transfer dictionary with virtually all the power of a direct machine translation system, in that much syntactic and semo-syntactic work can be performed in the course of lexical transfer.

Because of the production orientation, pre- and post-editing is greatly automated and assisted. While Bennett and Slocum speak of a DEC-20 environment, it should perhaps have been noted that the other locations in which pilot METAL programs exist handle these automated functions on a different, possibly more common, platform.

Discussion of the parser (a bottom-up strategy that services paths in parallel and maintains a variety of strong provisions for the prevention or early termination of unlikely paths) is clearly presented, in keeping with Slocum's strong understanding of parsing theory and thus its salient issues.

Discussion of the automatic speller is somewhat unexpected, in that the procedures for spelling checking and correction are quite well worked out and available on the smallest of word processors; the ability to attempt a spelling correction on the fly, without user intervention, is the most notable feature of the METAL speller.

This paper contains some of the most persuasive discussion of the evaluation metric for machine translation, and what such measurements mean as against abstract conceptions of system performance. Bennett and Slocum agree with the contention that there are critical unsolved problems in computational linguistics, but note that the assertion that these problems make machine translation systems unusable must surely be subject to empirical testing. Production systems survive as a direct consequence of their cost effectiveness, and only indirectly from their linguistic performance; thus success can be judged by the time of overall machine translation process versus time of overall human translation process. The data provided for the METAL

system (and indeed, data given in this volume for other systems) directly demonstrate the success of machine translation.

Makoto Nagao, Jun-ichi Tsujii, Jun-ichi Nakamura, "The Japanese Government project for machine translation"

The most notable part of the paper by Nagao *et al* deals with evaluation criteria, accuracy, and intelligibility, which appear at first to be so overlapping as to be useless as measurement dimensions. Presumably, accuracy could only be determined if a fairly complete intelligibility were present. It turns out, though, that the two criteria are in fact distinct enough to allow a meaningful evaluation of the success of a system's machine translation capability.

The Japanese Government project reported here was a four-year project to determine the feasibility of Japanese–English machine translation. The prototype system is a transfer-type, pre-edit-capable system that employs a significant amount of transformational power in its transfer lexicon.

The GRADE development environment provides a workbench for grammar writers, enabling creation of contextually driven and lexically driven grammar rules. The grammar writer in GRADE has the power to examine the consequences to ambiguity in the rules he/she writes, allowing a comparison of paths and rule weighting to avoid unnecessary ambiguity.

Much space is given to the power of the transfer lexicon to perform syntactic operations. This attribute of transfer systems is an important one to exploit, but there are perils for a system headed for production. First, attaching grammar rules to lexical entries is surely something that product end-users don't want to do, any more than software support people want them to. Secondly, there are consequences to modularity and maintenance whenever processes of a particular type are performed in two different places. Using the lexical rule power only in exceptional cases, and pre-coding these cases before releasing the software, seems to be the best employment.

This paper also has significant value in training newcomers to the machine translation world. There are, of course, a variety of contrastive problems that exist between any two languages. Numerous examples are provided of the natural contrastive issues and the computational complications of handling those issues. The clear presentation of many-to-many problems is very useful for understanding the nature of contrastive problems generally, beyond just English–Japanese.

Nagao *et al* present perhaps the best representation of experimental method in the volume. The researchers began with a corpus for translation, using half of the text for lexical and domain-specific phenomena, and tested on the other half. This allows a strong methodological support to experimental validity, while permitting pre-

test development to reduce problems of interpreting the test results.

Muriel Vasconcellos and Marjorie Léon, "SPANAM and ENGSPAN: Machine translation at the Pan American Health Organization"

The theme of this volume, that of systems and their relationship to production translation, is given a clear view from the production side by Vasconcellos and Léon. The Pan American Health Organization (PAHO) systems were conceived from the outset as production translation systems. The lessons learned from the operation of the direct-type SPANAM, and later the transfer-type ENGSPAN, led not only to the computational linguistic innovations that enhance accuracy and speed, but also the procedural and word-processing practices that are, ultimately, what makes these systems functional in the production translation workplace.

The translation activities in PAHO have been supported by MT since 1980, with the introduction of the Spanish-English system (SPANAM), followed in 1984 with the English-Spanish (ENGSPAN). Based on IBM mainframes from the beginning, the systems improve translation (including the comparative human/machine revision time) by two to three times. Additionally, components of the system support on-line human translation.

Several of the translation facts about PAHO, challenges that most of the current development systems have tried to avoid or have not confronted, have proven fortuitous. There is no one subject area domain in the translation work required by PAHO; nor is there any mechanism by which input language can be constrained, stylistically or syntactically. Thus descriptions of English and Spanish have to be generic and as complete as possible.

Perhaps the most significant points in this paper are the discussions of issues that the study of machine translation has avoided, because it appears not to be part of "interesting" linguistics. The "simple" issues of softcopy input and word-processing techniques and procedures in post-editing are, as anyone who has attempted production MT knows, consuming efforts that are almost thankless owing to their apparent lack of scientific interest. Thus it is encouraging to see discussions of these problems in print. Machine-readable input is particularly difficult to resolve within traditional mainframe operating system constraints. And the importance of compatible, tailorable word-processing capability can be seen immediately from an examination of the ALPAC report: most of the examples of unusable translation in that document would be considered readily usable today, the difference being that there were no modern text editors then. Discussions of these seemingly irrelevant areas are a strong contribution to the volume.

Pierre Isabelle and Laurent Bourbeau, "TAUM-AVIATION: its technical features and some experimental results"

Isabelle and Bourbeau rose to the challenge of reporting upon a project whose continuation had been cancelled, in this paper. The University of Montreal project had developed the highly successful METEO system for translating weather forecasts. The follow-on AVIATION effort, for translating aircraft maintenance manuals from English to French, was stopped in 1981, when it was determined that there was no immediately cost-effective production capability in sight.

The TAUM-AVIATION system is a heavily batch-oriented transfer system. There is a significant reliance upon the notion of "sublanguage" in the design, coupled with a higher level language independence. The authors allude to a theoretical model of human translation, which lends psychological validity to the transfer approach. A citation of this model would have been welcome; many readers who have the orientation of machine translation as a translation process would be vitally interested in such studies.

Committed to handling certain practical aspects of production translation, the TAUM-AVIATION project had a mechanism for maintaining formatting codes, so as to preserve format in the output. However, the authors contend that "fail-soft" techniques such as word-for-word or phrasal translations upon failure are worse for the translator than just outputting the original source language. This conclusion, while claiming to be oriented toward the user, is nevertheless the opposite of the conclusion reached by other projects reported in this volume. Lexical or phrasal translation can, it is true, be occasionally maddening to the posteditor. However, such outputs usually provide accurate terminology for the right translation, and are occasionally usable despite the computational failure. Further, these fail-soft outputs provide a means by which the translator can provide expertise in feedback to the programmer, enabling improvements more readily than if the output provides no specific clues about translation failure.

The transfer component of the TAUM-AVIATION system possesses the power to perform the conversion of argument structures to those required by target language predicates. The transfer system contains rules of transformational power, an ability that the authors claim has not received the treatment in the MT community that it deserves. While the value of transformationally-powered transfer algorithms is in fact covered extensively in both the Bennett *et al* and Nagao *et al* papers in this volume, the importance is perhaps worth repeating.

This is the only study in the volume that shows a machine translation system to be less cost-effective than human translation. TAUM has disbanded, and thus there is no easy way to evaluate the post-mortems. Yet

it is tempting to suggest that a different treatment in the user-access issues of lexical entry, word processing for post-editing, and fail-soft outputs might have affected the outcome favorably.

Jonathan Slocum, "A machine(-aided) translation bibliography"

The third paper in which Slocum was involved is a bibliography of machine translation, including machine-assisted translation and terminological studies. Perhaps the most up-to-date piece in the volume, this is another candidate for periodic updating, and constitutes by itself a sufficient justification for acquiring the volume.

Machine translation systems is a valuable collection of papers gathered with a view toward relating the research models with both commercial production needs and existing production systems. The criteria for success of a system-oriented project must, as Slocum maintains, focus on the comparison of system performance against human performance, taking into account the processes that both methods share. It is hoped that the perspectives of this volume will persist, whether in future collections, or in an updated version of this one.

John S. White has been involved in machine translation research and development since 1977, and was project manager of the METAL project from 1984 to 1986. His address is: Planning Research Corporation, 1500 Planning Research Drive, McLean, VA 22102.

NATURAL LANGUAGE PROCESSING

Hugh M. Noble

(Robert Gordon's Institute of Technology, Aberdeen)

Oxford: Blackwell Scientific Publications, 1988, xii + 240 pp.

(Artificial intelligence texts)

ISBN 0-632-01502-0, \$33.50 (sb)

[Editor's note: This book is reviewed twice.]

Reviewed by

Elena Pascaleva

Bulgarian Academy of Sciences

Useful survey books devoted to problems of natural language understanding and computational linguistics make use of several main approaches to the presentation of the material:

- Illustrating concrete research with a chosen formalism in order to achieve a gradual introduction to the subject matter (depth-first approach).
- Analysing a number of most representative contemporary methods of presentation and processing of linguistic knowledge in order to cover the whole range of achievements in the field (breadth-first approach).
- The two above-mentioned modes of description can be combined with the enumeration of problems that

are still unsolved but crucial for the progress of research.

Noble's book *Natural language processing* is a successful combination of all three approaches, each of them dominating in different parts of the book.

Part 1 presents a description of an extremely simple NLP system by means of the ATN formalism and the programming language POP-11 (partly Prolog, too). The result of the operation of the system—which later is described consistently in a highly limited microworld, adhering to the principle "from the simple to the complex"—is the parsing of basic English syntactic structures of the simple sentence. The presentation is a typical illustration of the first approach listed above. The author displays an admirable ability to introduce the reader to serious linguistic problems, each of them worthy of separate treatises (e.g., the formalized description of the English temporal systems) within a microworld consisting of three points, connected by a line.

Part 2 is an enquiry into the semantics of NLP systems (making use of the second approach). Several alternative means, already successfully adopted in AI systems, are proposed. These proposed means are aimed at overcoming the incompleteness of the linguistic description proposed in Part 1. For the purpose of successively acquainting the reader with these means, the following model of presentation is used: every new approach suggests means of overcoming the failures of the previous one. The order is thus: case grammar, frames, scripts and plans, and conceptual dependencies. At this stage, the illustrations do not contain concrete programs, but rather only examples for the operation of the basic formalisms.

Part 3 (where application aspects have definitely moved to the background) is an attempt to analyze various aspects of real NLU, i.e., natural language understanding by people. Raising the slogan to work out a theory of semantic presentation, already unlimited by concrete microworlds, and at the same time proposing a formalism of his own for such a presentation, the author leaves the field of AI systems and enters the sphere of the cognitive and pragmatic. A large number of problems are touched upon: perception, motivation, causation, models of communication, knowledge acquisition, metasemantic problems, cognitive aspects of truth, falsehood, negation, quantification, and others. Since only 73 pages are devoted to this huge amount of problems, the author obviously only manages to draw our attention to summits that are yet to be conquered, not only through CL and AI, but also through general linguistics. The height of these summits clearly demonstrates the distance that is yet to be covered in order to achieve real NLU modeling.

Though extremely concise, the presentation in this part of the book, with the impressive range of problems discussed brought together by a unified philosophical