

Varile, G. B. & A. Zampolli (eds) (1997), *Survey of the State of the Art in Human Language Technology. Linguistica Computazionale*, vol. XII-XIII. Giardini Editori e Stampatori in Pisa, Pisa, Italy. Cambridge University Press ISBN 0-521-59277-1. Giardini Editori. ISBN 88-427-0018-5. 513pp.

Zarco Tejada, María Ángeles

*Universidad de Cádiz, Departamento de Filología Francesa e Inglesa,
Facultad de Filosofía y Letras, C/ Bartolomé Llompart s/n, 11003 Cádiz,
Tfno.: 956-245087, Fax: 956-220444, e-mail. azarco@galeon.uca.es*

(Recibido Enero 1998; aceptado Marzo 1998)

BIBLID [1133-682X (1997-1998) 5-6, 493-498.]

This volume offers a survey of the major results and on-going research issues in human language technology. It gives a comprehensive overview of the state of the art in spoken and written language technology: the main areas of work, capabilities and limitations of current technology and the technical challenges that must be overcome to realize human computer interaction using natural communication skills.

It is organized according to thirteen chapters written by 97 different authors, that show the main scientific and technical areas of this field. That is to say, spoken-written language input-output; language analysis, generation and discourse; document processing; multilinguality; application of mathematical methods; language resources and a final evaluation on the different fields. A consistent structure is achieved all along the book. Each of the thirteen chapters consists of an introductory overview of the area under analysis, a number of sections dealing with the state of the art for a given sub-area, a 'future directions' section where key research fields are considered and a final specialized bibliography.

The study of human language technology is a multidisciplinary enterprise requiring expertise in areas of linguistics, psychology, engineering and computer science. This book is an excellent piece of work not only for the content -a very good current description of the state of the art-, but for being written by the most predominant experts on each specialized sub-area. Among others: Ronald Cole, Victor Zue, Joseph Mariani, Annie Zaenen, Fernando Pereira, Ted Briscoe, Stephen G. Pulman, Hans Uszkoreit, Martin Kay, Antonio Zampolli, Ken Church, Ralph Grishman and Nicoletta Calzolari.

The first chapter of this survey deals with spoken language input technologies. In the near future, interactive networks will provide easy access to information and services but, today, such networks are limited to people who can read and have access to computers. Advances in human language technologies are needed to enable the average citizen to communicate with networks using natural communication skills. Much effort has been done on the development of spoken language interfaces and our authors describe all the stages needed that will facilitate such an interface. Thus, the basic issues of speech recognition are described: signal representation, acoustic modeling (the most widespread technique being hidden Markov modeling), speaker recognition technologies and language identification. The natural language component, working in concert with the recognizer, produces a meaning representation. This process is explained in the final section of the chapter where the integration of speech recognition and natural language processing techniques is discussed.

Parallel on the organization of the problems described but differentiated by the type of channel language input takes, the second chapter starts paying attention to the problem of recognition of written language input. The nature of written language is described, how written language is transduced into electronic data and the nature of written language recognition algorithms. Written language recognition transforms language (printed text as well as handwriting) into its symbolic representation. The central tasks are character recognition and word recognition. The two basic techniques for character recognition, Optical Character Recognition (OCR) in the case of a print image, and Intelligent Character Recognition (ICR) in the case of handprint, open up the Recognition sub-section. Handwriting recognition aims at designing systems able to understand personal encoding of natural language. Problems such as ambiguity derived from missing letters, or poorly writing and strategies to solve these problems such as the holistic approach and the analytical approach are addressed.

Word recognition is handled by language models, being the lexicon of words, determined by linguistic constraints, the most important one. Finally, a further step in the recognition of written language is achieved in the sub-section on document analysis. This is concerned with the global issues involved in the organization of the document structure.

There is a final exposition of the state of the art and useful techniques in on-line handwriting recognition (handwriting recorded with a digitizer in contrast with off-line handwriting recognition which is optically scanned). As it is stated, on-line handwriting recognition has borrowed a lot of techniques from OCR and speech recognition.

Chapter three focuses on the analysis of language within a natural language processing perspective. Under this view, different issues are considered, e.g. sub-sentential/sentential processing, grammar formalisms, the lexicon, semantic theories and computational aspects.

By sub-sentential processing our authors mean morphological analysis, morphological disambiguation and shallow parsing. Though, as it is affirmed, word-form analysis and generation is no longer an academic research topic, the main problems and approaches to it are enumerated, that is to say, morphological alternations and morphotactics as main problems, and the cut-and-paste method and the two-level model as morphological approaches. A good survey of the different models designed to handle morphological disambiguation and shallow parsing is given.

Whereas we find a very brief exposition of grammar formalisms, -only constraint-based grammars or unification grammars are considered, and a lexicon, only for constraint-based grammars as well-, semantics is much more detailed. First order predicate calculus (FOPC) is the topic of the initial sub-section on semantics followed by a more powerful logic, Montague intensional logic. Montague-derived approaches such as dynamic semantics, situation theory and property theory are also referred. Finally, computational issues (precision, coverage, search procedure) related to different types of sentence grammars (linguistic grammars, task-oriented grammars and data-oriented grammars) are discussed.

The subject under study after language analysis is language generation. Chapter four deals with written language generation whereas chapter five is about spoken output technologies.

As regards the former, and before getting into it, an explanation of the tasks and process of generation as well as a survey of significant problems is provided. Three types of generator task are identified: text planning, sentence planning and surface realization. On the other hand, generator processes are classified according to the range of sophistication and expressive power. Starting with inflexible canned methods and ending with maximally flexible feature combination methods, the following systems are described: canned text systems, template systems, phrase-based systems and feature-based systems. Lexical selection, sentence planning, discourse structure, domain modeling and generation choice criteria are enumerated as the remaining significant problems for any general-purpose sentence generator.

Generation is subdivided into two sub-sections, syntactic generation and deep generation, due to their specific goals: the former the task of finding a string for an input logical form, the latter, the production of specifications of linguistic abstraction to drive surface generators.

The chapter on spoken output technologies gives us a general exposition of the main research areas and problems of spoken language generation. Even though, a clear distinction is made between synthetic speech generation and spoken language generation on the bases of the kind of input (Text-to-Speech processing models not only human speech generation but also text processing), the problems discussed are fairly similar: prosody control, speaker characteristics control, intonation, intelligibility, naturalness, etc.

Once the levels of analysis and generation have been discussed, The Survey takes us to consider a further level of relations, the one beyond sentence interpretation, Discourse and Dialogue. The major lines of research on determining what information carries the text fall into two categories: informational and intentional. Discourse and dialogue are considered independently. We are given the most important issues of research in both areas as well as actual systems and research groups dealing with context, but we are given strong emphasis on a common central problem, the lack of a well-defined theory of dialogue where an integration of semantics and pragmatics were achieved. Current approaches to discourse and dialogue from an artificial intelligence and computational linguistics point of view are based on the following four theories of discourse: Hobbs (1985), Grosz and Sidner (1986), Mann and Thompson (1987) and McKeown (1985).

Another area of research covered by this Survey refers to the work that surrounds the use of documents, Document Processing. By bringing technology to the process of producing and using documents, significant productivity enhancements have been achieved. Thus, Internet access and sophisticated search algorithms driven against massive amount of data are having a major effect on document retrieval. But such massive amount of information cannot be digested as such, it needs being analyzed and structured, it needs being interpreted. Much of the key information in interpreting texts in these applications comes from discourse units. The increasing volume of machine-readable text has stimulated the interest for automatic summarization. This technique is described as having two major phases: *interpretation* of the source text into an overall *source meaning representation*, and *generation* of the summary by formation of the *summary representation* using the source one and subsequent synthesis of the summary text. As a consequence of the application of technology to document processing and because of efforts on simplification of the problems, Controlled Languages (a form of language with special restrictions on grammar, style, and vocabulary usage) have been developed to improve the consistency, readability, translatability and retrievability of information. As an example, the best known recent controlled language is AECMA Simplified English, adopted by the aerospace industry.

Chapter eight deals with tasks that involve the use of more than one natural language. Among all, translation, language identification, and the multilingual dimension of storing, retrieving information, and processing of spoken material are reviewed. As regards translation, a survey of Machine Translation, Human-Aided Machine Translation, and Machine-Aided Human Translation is given, being old and issues of today on this field the distinction between interlingual and transfer approaches, the importance of linguistic matters, and the type of user the system is designed for. Finally, multilinguality is considered within a spoken form. Research activities in this area come from spoken language identification, multilingual speech recognition and understanding,

and speech translation.

The following two chapters give a survey of more technical applications of technology to human communication. Thus, activities in image, voice and tactile interfaces for human-machine communication are commented with a final reference to projects, HuMaNet system, that integrate multiple modalities. Such sort of systems have important positive consequences on a medical side, whereas systems for amusement are required as well. Among modality integration types, speech and gesture interaction and facial movement and speech recognition and synthesis are referred. In fact, transformations from images to text and from text to images are a great challenge for computer systems.

On the other hand, the chapter dedicated to Transmission and Storage of information deals with two linked areas of speech processing: coding and enhancement. In the speech coding research field, coding algorithms seek to minimize the bit of rate in the digital representation of a signal without loss of signal quality, whereas speech enhancement has focused on the suppression of background noise.

A variety of mathematical methods have been employed in human language technology. This chapter shows an overview of the different approaches. Mathematical methods applied to syntax, phonology and morphology helped to describe structured sets of strings. Using mathematical tools, unification grammar is an example of a grammar that represents linguistic entities as a set of complex features. On the other hand, in semantics many varieties of higher order predicate logic have been developed. Finally, many researchers have turned to statistical data-driven methods for designing language technology applications. A special area developed during the last few years, in conjunction with statistical methods, is the utilization of optimization (defined as the process of finding a best solution) techniques for spoken and written language processing (Digital Signal Processing). Among low-level processing methods for natural language it is noteworthy the renaissance of finite-state processing techniques.

Before arriving at the last chapter, the one on evaluation, the survey on language resources gives us a view on the sets of language data in machine readable form. These linguistic resources are written and spoken corpora, lexical databases, grammars and terminologies and are used in building, improving or evaluating natural language.

A controversial problem, underlined here, is whether linguistic resources must be theory-neutral, so that these can be reusable and multifunctional. The major action lines on the organization level are given by the most important associations on Computational Linguistics, ACH (Association for Computing in the Humanities), ALLC (Association of Literary and Linguistic Computing) and ACL (Association for Computational Linguistics) with their Text Encoding Initiative. Apart from very accurate information on current language resources and a description of future directions, a very helpful large list of addresses for contact information on corpora is given. We find there

not only the list of any project mentioned before, but the address, contact person and e-mail.

Finally, the last chapter introduces us to the notion of evaluation in speech and natural language processing. It plays an important role for developers, for system integrators and for consumers, since such sort of analysis tells if the system is improving, which approaches to follow and which system will best meet a specific set of needs. In this chapter three kinds of evaluation are broadly distinguished: adequacy evaluation (determination of the fitness of a system for a purpose), diagnostic evaluation (production of a system performance profile with respect to some taxonimization of the space of possible inputs) and performance evaluation (measurement of system performance in one or more specific areas). Evaluation is displayed for text analysis, machine translation and translation tools, natural language parsers, users, speech input and synthesis, usability, speech communication quality and character recognition.