The Penman Natural Language Project

Information Sciences Institute
of the University of Southern California
4676 Admiralty Way
Marina del Rey, CA 90292-6695
Tel: 213-822-1511

Objectives

The development of useful and theoretically motivated computational resources for natural language input and output. This work may in future take the form of the development of a multilingual knowledge-based machine-aided translation system, together with the Center for Machine Translation at CMU and the Computing Research Laboratory at New Mexico State University.

Research Approach and Endeavors

Following the idea that natural language cannot be adequately handled by addressing only syntax and semantics, the project seeks to draw upon a wider theoretical basis, primarily that of Systemic-Functional Linguistics. The Penman sentence generation system contains one of the largest-coverage computational grammars of English, and has been distributed to approximately two dozen research and academic sites throughout the world. In collaboration with the Linguistics Department of the University of Sydney and through them of a worldwide network of systemic linguists, grammar development is an ongoing process.

In addition to sentence generation, the project is conducting research on parsing. A parser based primarily on classification with respect to a subsumptively-organized concept lattice is currently under development. By simultaneously accessing syntactic and semantic knowledge, this parser promises an improvement of efficiency and organizational clarity over many older methods, among others over traditional unification, with which it shares many aspects.

A third major area of interest is text planning. In collaboration with other researchers at ISI and elsewhere, project members are working on a method of planning multisentence and eventually multiparagraph texts using interclausal relations tied to discourse intentions.

In collaboration with CMT and the CRL, the parser and generator and various other auxiliary knowledge sources are being assembled into a machine-aided translation system. Various collaborators, both in the USA and in Europe, are conducting related research; for example, the KOMET project at IPSI in West Germany are building a German grammar for Penman.

Information and Contact Addresses

The Penman project has a substantial publication record, publishing in both the computational (ACL, Coling, AAAI, IJCAI, etc.) and linguistics (Systemics Congress, Text, IPRA, etc.) arenas. The Penman Upper Model and sentence generator are available for research and educational purposes (systems run on TI Explorer and Symbolics Lisp machines, MacIntosh-IIs, and Sun 3s and 4s). Please contact Dr. Eduard Hovy at the above address or at HOVY@ISI.EDU.