ARTIFICIAL INTELLIGENCE IN ENGINEERING

October 21–23, 1985

George Washington University
Marvin Center
21st & H Streets, N.W.
Washington, D.C. 20052

RULE-BASED DATA BRIDGES
PROVIDE EXPERT SYSTEMS AND APPLICATIONS
WITH CRITICAL DATA SOURCES

Richard Roth

ABSTRACT

A major stumbling block to use of expert systems and many new computerized applications in the business world is availability of current information. In many cases the information required originates on another computer but is only available in a form intended for human consumption, usually a report. The nature and variety of these reports makes producing machine-understandable data in to a task requiring AI techniques. The context of the information is significantly different from that found in natural language processing or visual pattern recognition, and requires development of new techniques.

A system has been built that embodies a rule-based parser to process a printed type report into an internal format data file. Using pre-defined format rules and output goals, the parser is capable of producing a standard application defined data file from a printed report file. The rules are currently developed via a screen-oriented editor with the user highlighting the required format details and needed data. The project is now extending the automatic recognition rules to minimize the need for user guidance during the rule building stage.

The possibilities for this system point in two directions: providing a source of data for expert systems, and the broader aspect of general data bridging. A "data bridge" allows tying two computer systems together via information not intended for inter-machine transfer. Since most real applications produce information in specialized formats, the ability to make this information available in a different format without major investment in resources, opens the potential for more useful computer systems. This impact can be felt most significantly when combined with the rapidly growing power and dropping cost of personal computers. The main problem with larger computer systems was not their computing power, but the limited human resources to properly program them. Just as other AI techniques allow the computer to interact in a human-oriented world, so too, a data bridge makes the information produced by existing computer systems available for wider use than intended by the original system. Most importantly, these techniques allow gaining the maximum benefit from the processing done by the existing system without changes to the existing system.

The system described here runs on an IBM personal computer and directly processes reports from larger systems, available by standard communications methods. It produces data files that may be used by existing application tools such as the Lotus 1-2-3 spreadsheet or the Dbase database system. The original application described here, which handles financial data from bank statements, is only the tip of the iceberg of potential uses for the system.