AI Magazine Volume 7 Number 5 (1986) (© AAAI)

RESEARCH IN PROGRESS

Ernest Davis and Ralph Grishman

AI Research in Progress at the Courant Institute, New York University

The AI Lab

In the following five subsections, we present a brief discussion of research in natural language understanding, three-dimensional model-based vision, domain models for commonsense reasoning, a text-fragment approach to support knowledge exploration, and finally, learning and expert systems.

The Proteus Project: Natural Language Understanding

The central goal of the Prototype Text Understanding System (PROTEUS) Project is the development of robust systems for natural language text understanding. Although the techniques being studied should be widely applicable, we are specifically developing a system to understand paragraph-length messages about equipment failures, with the aim of summarizing each failure and assessing its impact.

Several laboratory prototypes have been constructed for similar text analysis tasks, but none are reliable enough for practical application. We aim to improve on these earlier systems through a combination of two techniques: the use of detailed domain knowledge to verify and complete our linguistic analyses and the use of "forgiving" algorithms that can obtain an analysis even if some syntactic or semantic constraint is violated.

To guide the development of our system, we selected a corpus of messages describing the failure of one particular piece of equipment, a starting air compressor. We developed a grammar and implemented a parser for these messages. Recently, we began developing a detailed model of the structure and function of this piece of equipment. This model plays several roles in the system: It provides information for syntactic analysis, in particular for the analysis of the long compound nominals that occur frequently in such texts; it is used in establishing the causal relationships among the events mentioned in the message; and through display routines associated with parts of the model, it provides visual feedback of the system's understanding of a message. We are also experimenting with diagnostics to be issued when the input violates a semantic constraint, diagnostics that should be understandable to users without any specialized linguistic training (Grishman, Ksiezyk, and Nhan 1986).

This research is being conducted in conjunction with a group at System Development Corp. (Paoli, Pennsylvania), with each group responsible for certain aspects of system design. Our groups are jointly responsible for integration of the next-generation text-processing system as part of the Defense Advanced Research Projects Agency (DARPA) Strategic Computing Program (Grishman and Hirschman 1986).

In addition to the message-processing system, we maintain a small question-answering system that answers simple English queries about a student transcript database This system is used for teaching and as a preliminary test bed for some of our linguistic analysis techniques.

Participants: Ralph Grishman (faculty); Tomasz Ksiezyk, Ngo Thank Nhan, Michael Moore, and John Sterling

Abstract The AI lab at the Courant Institute at New York University (NYU) is pursuing many different areas of artificial intelligence (AI), including natural language processing, vision, commonsense reasoning, information structuring, learning, and expert systems. Other groups in the Computer Science Department are studying such AI-related areas as text analysis, parallel Lisp and Prolog, robotics, low-level vision, and evidence theory

Ernest Davis is an assistant professor of computer science and Ralph Grishman is professor of computer science and chairman of the Computer Science Department, Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012

(research staff); and Dimitri Turchin, Leo Joskowicz, Ping Peng, and Mahesh Chitrao (students).

Three-Dimensional Model-Based Vision

We have been developing the theory and implementation of computer vision systems for recognizing known three-dimensional (3-D) objects from unknown viewpoints in single two-dimensional (2-D) images. This generalized "bin of parts" problem requires a computer to recognize objects from any viewpoint even when they are partially occluded and viewed in a cluttered background. In addition to its practical applications in robotics, the work is aimed at providing a theory for several basic aspects of human visual recognition. We have placed particular emphasis on developing robust solutions for recognition, which make full use of the prior knowledge regarding an object's appearance to reliably interpret partial and locally ambiguous data from the image.

An important constraint for achieving robust recognition is the viewpoint-consistency constraint, which requires that the locations of all object features in an image must be consistent with projection from a single viewpoint This constraint has been applied by developing algorithms for solving for the 3-D position and orientation of an object directly from 2-D image measurements. The mathematical method used is based on an iterative Newton-Raphson solution for the best viewpoint parameters that "project" the object model onto the locations of matched image features. Once a few initial matches have been formed, it is possible to make quantitative predictions for the exact locations of further model features in the image.

A second part of the research concerns the problem of perceptual organization (Lowe 1985a). Human vision is able to spontaneously detect many different types of significant groupings of image elements. For example, people immediately notice significant instances of parallelism, collinearity, proximity, or symmetry in an otherwise random set of image features. This capability is missing from almost all computer vision systems. Because these image groupings reflect viewpoint-invariant aspects of a 3-D scene, they are ideal structures for bridging the gap between the 2-D image and the 3-D model.

Much of this research has been combined into a functioning system for visual recognition, named Spatial Correspondence, Evidential Reasoning, and Perceptual Orgnanization (SCERPO) (Lowe 1985b, 1986). This system was one of the first to demonstrate the recognition of 3-D objects in single images taken from arbitrary viewpoints.

Participant: David Lowe (faculty).

Domain Models for Commonsense Reasoning

We are studying a number of domains of commonsense reasoning and trying to construct formal models to describe them. One area we are studying intensely is the relation between spatial and physical reasoning. Our aim is to design a system capable of reasoning about the physical properties of simple rigid objects, such as boxes, hangers, hooks, and the like, and relating their physical properties to their shapes. To date, we have concentrated entirely on the representational issues of defining an ontology and a formal language sufficient to describe shapes, positions, events, actions, and physical constraints. Substantial progress has been made, particularly in defining these concepts ontologically and in relating topological shape descriptions to physical properties of objects (Davis 1984a, 1984b). We are now chiefly involved in developing a language of shape description suitable for physical inferences. The requirements that physical reasoning makes of a shape description language turn out to be different (in significant and interesting ways) from the requirements made by vision systems or purely geometrical reasoning systems.

Davis is also collaborating with Sanjaya Addanki of International Business Machines, Thomas J. Watson Research Center laboratories on the Physical Reasoning by the Organization and Modification of Prototypes (PROMPT) system, which uses multiple physical theories at various levels of depth to analyze complex physical systems. A prototype version of PROMPT was implemented for simple block world problems (Addanki and Davis 1985).

Finally, we have been studying the quantitative reasoning necessary for spatial and physical reasoning. In particular, we examined the power of constraint propagation as an inference technique for real-valued quantities (Davis 1985).

Leora Morgenstern (1986) has been researching the logic of knowledge needed to construct plans that involve "finding something out" as a substep. She has been using a syntactic approach in which the knowledge operator is a first-order predicate that takes as its arguments an agent and a quoted string which represents a proposition. Such a language is extremely flexible in the kinds of propositions about knowledge and ignorance that it can express; Morgenstern has successfully exploited this flexibility to produce a powerful axiomatic system relating knowledge to planning. The price of this flexibility is that it is possible to construct self-referential sentences, such as, "I know this sentence is false." However, following an approach due to Saul Kripke (1975), Morgenstern has shown that the existence of these sentences does not necessarily lead to a contradiction in the logic.

Participants: Ernest Davis (faculty) and Leora Morgenstern, Leo Joskowicz, and Alex Botta (students).

Netbook: A Text-Fragment Approach to Support Knowledge Exploration

Knowledge exploration is the activity of finding out what others have thought and written about. The goal of the Netbook Project is to help knowledge explorers find appropriate text fragments or ordered sets of text fragments from precanned text. The approach is to use a query system similar to that found in databases, augmented with domain-independent heuristics.

In our running prototype, explorers navigate through a type-function structure (analogous to frame-slot, object-handle, and relation-attribute structures). Associated with each attribute is a partial order that hierarchically orders the values of the attribute's domain; for example, Stockholm is a subvalue of Sweden in the location domain.

This structure has five important features. First, it permits explorers to descend the hierarchy associated with one attribute independently of the hierarchy associated with another. For example, if two attributes of events are time and location, one can specify the time to be in the 1890s and the place to be southern Europe or (further down the hierarchy) France or (again, further down the hierarchy) Paris. Second, the system permits explorers to choose two disjoint alternatives in a hierarchy, for example, France or England. This feature is directed against the complaint that explorers are forced to make choices in menu systems without knowing the difference between the choices. Third, the system permits the attributes to behave as two attributes when choices become orthogonal. For example, a user exploring the functions of a text editor can be interested in movement commands as opposed to modifying commands. Under movement, the attribute can split into the choices of direction (forward or backward) and length (line, paragraph, and so on). Fourth, the structure encourages further exploration to topics that are related by the hierarchy. In the editor example, a person who learns how to move a line down might then be curious about how to move a paragraph up. Fifth, the hierarchy helps the system capture generalizations such as, "in vi, deletion and movement commands are closely related."

The "intelligence" of the system consists of three abilities: to group text into clusters based on information-retrieval clustering techniques to present fragments in accordance with prerequisite relationships and to generate examples based on a grammatical characterization of a knowledge domain where appropriate (for example, when the domain is a procedure or a language) (Shasha 1985, 1986).

Participants: Dennis Shasha (faculty) and Leah Beckman and Jose Perez-Carballo (students).

Learning and Expert Systems

Our work in learning has been concerned with learning heuristics from a trainer. Specifically, the trainer supervises the behavior of a potentially nondeterministic program by specifying the appropriate actions at decision points. The trainer also specifies the goal structure being used. The program then attempts to integrate this information into its set of heuristic decision rules, using generalization when possible. This work was done in conjunction with Paul Benjamin, now at Phillips Research in New York (Benjamin 1984).

We are also in the process of developing an expert system for use in hospital intensive care units. The system takes real-time inputs from various transducers and uses medical knowledge to detect abnormal or dangerous conditions. The area of concentration for the initial implementation is post-cardiac surgery; however, the system is prototype driven and appears to extend in a fairly linear manner. This work is being done with Steven Henkind in a joint program between NYU and Mount Sinai Medical School.

We have also been looking at the optimization of expert systems for use in real-time control systems.

Participants: Malcolm Harrison (faculty) and Steven Henkind (student).

AI Research: Other NYU Groups

Several other research groups at NYU are studying issues closely related to AI.

Linguistic String Project

The NYU Linguistic String Project, led by Professor Naomi Sager, has been active in natural language-processing research for 20 years. Members of the project were pioneers in the development of broad-coverage grammars for text analysis (Sager 1981) and the study of the information structures of sublanguages (the language used in restricted scientific and technical domains) (Grishman and Kittredge 1986). These techniques are being applied to the analysis of medical texts and patient records and to the automatic mapping of language-borne information into a database for subsequent retrieval.

Participants: Naomi Sager and Paul Mattick Jr. (faculty); Emile Chi, Judith Clifford, Carol Friedman, and Stephen Johnson (research staff); and Joan Chen, Christina Borges, and Catherine McLeod (students).

Parallel Programming Languages for AI

The NYU Ultracomputer Project is a major effort in the development of an architecture, operating system, and support software for a large general-purpose shared-memory parallel processor. A small hardware prototype and a parallel UNIX-based operating system are currently operational. As part of this effort, we have become involved in the study, design, and implementation of AI programming languages for the ultracomputer. Initially, our work concentrated on implementations of parallel Lisp and Prolog.

Participants: Malcolm Harrison (faculty) and Isaac Dimitrovsky, Chin Woo Yoo, Nick Markantonatos, and Georgios Papadopoulos (students).

Research in Robotics at NYU

Robotics research at NYU focuses on the mathematical and software foundations of the field. Current areas of emphasis are (1) algorithms for the analysis of sensory (especially visual) data; (2) software and control techniques for dextrous motion of multifingered hands; and (3) theoretical, algorithmic, and geometric issues in robotics, especially motion planning. These topics combine classical applied mathematics, algorithm design and analysis, and AI research.

Work in the area of robot sensing has concentrated on the mathematical aspects of visual-sensor algorithms, particularly the use of structured light and the study of 3-D shape- and object-matching techniques. We have developed sensors that make 3-D range data available and subsequently used this form of data (and simpler 2-D image data) to locate and characterize the objects present in an image with their orientations. Because a 3-D continuum of orientations is possible and because the 3-D data provided by a sensor can be corrupted by noise, the matching problem that arises is far from trivial. We have developed new and effective approaches to the problem that utilize local geometric invariants to ''filter'' plausible match candidates from large vocabularies of possible objects, giving us the ability to recognize objects in a time independent of the size of the vocabulary from which the objects are drawn.

There is ongoing research into the control of dextrous multifingered hands. This area includes work on software able to accomplish a variety of dextrous motions and aims at robots that can adjust smoothly and simply to the shapes and physical behavior of delicate 3-D bodies. Sophisticated multidimensional feedback control methods are needed and represent one component of our work. This work has both an experimental and a theoretical side. On the experimental side, we have constructed an experimental nonanthropomorphic four-fingered hand, with the hardware and control software needed to interface it to our VAX systems and to move the fingers at acceptable speed. Software that allows a single finger to track edges of unknown shapes has been constructed, and we are now in the process of extending this experimental work to involve two, three, and four fingers. On the theoretical side, several simulations of nonlinear control heuristics suitable for multidimensional nonlinear control have been carried out. These simulations include robot control and control in the well-known robot ''crank-turning'' problem.

A third area of research consists of theoretical studies of geometric issues in robotics. Our work in this area concentrates on issues connected with automatic planning of paths between specified points, avoiding obstacles whose geometry and position are known. Known methods for solving these problems (many developed by our group) are still infeasible computationally. Our aim is, therefore, to find much faster algorithms that work by a principle of successive approximation, by exploitation of the special properties of some appropriately limited class of surfaces, or by a relation of geometric problems set in high-dimensional spaces to versions of the same problems projected into lower-dimensional space. Several such algorithms have already been constructed for simple special systems of bodies (Schwartz, Sharir, and Hopcroft 1986).

Participants: Robert Hummel, David Lowe, Bhubaneswar Mishra, Jacob Schwartz, and Micha Sharir (faculty).

Representation of Images

Robert Hummel is conducting a research project on low and intermediate levels of image representation. Funded by the Office of Naval Research, the project is interdisciplinary and is collaborative research with Professor Michael Landy in the Department of Psychology. The researchers and their students are interested in the first few layers of transformations that image data undergo in human visual processing. By understanding and mimicking these representations, they hope to improve methods for extracting information and making inferences from digital imagery.

The principal research ideas center around multiresolution representations and relaxation labeling methods for knowledge aggregation. In particular, the Laplacian pyramid structure as studied by Burt and Adelson (1983) seems to hold great potential as an intermediate representation for image analysis. This structure decomposes an image into its components at different scales: textural components at a fine scale and edge components and regions at coarser scales. By discarding information that psychophysical experiments tell us is unimportant to the human visual-recognition process, a good representation permits easy access to the features necessary to perform scene analysis.

Among the many possible methods of representing knowledge acquired from an image, the researchers are particularly interested in iterative methods that modify beliefs in propositions as evidence is accumulated. Models must be able to handle conflicting evidence, that is, to interpret objects in an image in the presence of noise and ambiguity. Methods for extracting and accumulating evidence include the Dempster-Shafer theory of evidence (Shafer 1976), stochastic relaxation methods, and relaxation labeling processes (Hummel and Zucker 1983).

Participant: Robert Hummel (faculty).

References

Addanki, S , and Davis, E 1985 A Representation for Complex Physical Domains In *Proceedings of the Ninth International Joint Conference for Artificial Intelligence*, 443-446

Benjamin, P 1984 Extraction and Generalization of Expert Advice. Ph D. diss , Computer Science Dept., New York Univ

Burt, P , and Adelson, E. 1983 The Laplacian Pyramid as a Compact Image Code In IEEE Transactions on Communications COM-31(4): 532-540

Davis, E 1985 Constraint Propagation on Real-Valued Quantities, Technical Report 189, Computer Science Dept , New York Univ

Davis, E. 1984a. An Ontology of Physical Actions, Technical Report 123, Computer Science Dept., New York Univ.

Davis, E 1984b Shape and Function of Solid Objects: Some Examples, Technical Report 137, Computer Science Dept , New York Univ

Grishman, R , and Hirschman, L 1986. PROTEUS and PUNDIT: Research in Text Understanding PROTEUS Project Memorandum 1

Grishman, R , and Kittredge, R , eds 1986 *Analyzing Language in Restricted Domains.* Hillsdale, N J : Lawrence Erlbaum

Grishman, R ; Ksiezyk, T ; and Nhan, N T 1986 Model-Based Analysis of Messages About Equipment. PROTEUS Project Memorandum 2

Hummel, R A., and Zucker, S W 1983 On the Foundations of Relaxation Labeling Processes In IEEE Transactions on Pattern Analysis and Machine Intelligence 5(3): 267-287

Kripke, S. 1975 Outline of a Theory of Truth *Journal of Philosophy* 72: 690-716

Ksiezyk, T , and Grishman, R 1986 An Equipment Model and Its Role in the Interpretation of Nominal Compounds PROTEUS Project Memorandum 3

Lowe, D G 1986. Three-Dimensional Object Recognition from Single Two-Dimensional Images: Technical Report 62, Courant Institute, New York Univ

Lowe, D G 1985a Visual Recognition from Spatial Correspondence and Perceptual Organization, In Proceedings of the Ninth International Joint Conference for Artificial Intelligence, 953-959

Lowe, D G 1985b *Perceptual Organization and Visual Recognition* Boston: Kluwer Academic

Morgenstern, L 1986 A First-Order Theory of Planning, Knowledge, and Action In Proceedings of the Conference on Theoretical Aspects of Reasoning about Knowledge, 99-114 Los Altos, Calif : Morgan Kaufmann

Sager, N 1981 *Natural Language Information Processing A Computer Grammar of English and Its Applications* Reading, Mass.: Addison-Wesley

Schwartz, J T.; Sharir, M.; and Hopcroft, J 1986 *Planning, Geometry and Complexity of Robot Motion Norwood, N.J.: Ablex*

Shafer, G. 1976 *A Mathematical Theory of Evidence* Princeton, N.J : Princeton University Press

Shasha, D 1986. When Does NonLinear Text Help? In Proceedings of the First International Conference on Expert Database Systems, ed. L Kerschberg, 109-120.

Shasha, D 1985 Netbook—A Data Model to Support Knowledge Exploration. Proceedings of the Eleventh International Conference on Very Large Data Bases, eds. A Pirotte, and Y Vassiliou, 418-425