January 1970

THE SRI INTELLIGENT AUTOMATON PROGRAM

by

John H. Munson

To be published in the Proceedings of the First National Symposium on Industrial Robots, Chicago, Illinois, April 2-3, 1970.

Artificial Intelligence Group

Technical Note 19

SRI Project 8259

The research reported here was supported by the Advanced Research Projects Agency and the National Aeronautics and Space Administration under Contract NAS12-2221.

THE SRI INTELLIGENT AUTOMATON PROGRAM

by

John H. Munson
Stanford Research Institute
Menlo Park, California

Introduction

A research group at Stanford Research Institute (SRI) is experimenting with an intelligent automaton, or robot, in order to study the fundamental cognitive processes that will someday enable a machine to perform useful tasks autonomously in a real environment.[1-4*†] We are not attempting to build a human-like robot: Although we take note of the increasing dialogue between physiologists and engineers, our approach is drawn more from computer science and engineering than from bionics and physiology. Our goal is to explore the basic functions that increasingly will be demanded in the computer control of mobile machines interacting with complex environments. These include sensory input and perception (especially vision), pattern recognition, natural-language conversation, goal structuring, cognitive model building, and problem solving. These topics span that area of present-day computer science that has come to be called "artificial intelligence" or "machine intelligence." Our motivations include both that of "pure" research (understanding intelligent systems) and of "applied" research (developing useful computer-controlled systems).

---

[*] References are listed at the end of this paper.

[†] This work is being sponsored by the Advanced Research Projects Agency and the National Aeronautics and Space Administration under Contract NAS12-2221.

The speaker will present a 20-minute film that introduces the SRI robot, shows it performing some simple tasks, and outlines the structure of the program in the computer that controls the robot.

The Vehicle and Its Environment

The SRI robot system consists of two parts, computer and mobile vehicle, connected by radio. The robot vehicle itself is shown in Figure 1. It is propelled by two stepping motors that independently drive wheels on either side of the vehicle. It carries a vidicon television camera and optical range-finder in a movable "head." Control logic on board the vehicle routes commands from the computer to the appropriate action sites on the vehicle. In addition to the drive motors, there are motors to control the camera focus and iris settings and the tilt and pan angles of the head. Other computer commands arm or disarm interrupt logic, control power switches, and request readings of the status of various registers on the vehicle. Besides the television camera and range-finder sensors, several "cat-whisker" touch sensors are attached to the vehicle's perimeter. These touch sensors enable the vehicle to know when it bumps into something. Commands from the computer to the vehicle and information from the vehicle to the computer are sent over two special radio links, one for narrow-band telemetering and one for transmission of the TV video from the vehicle to the computer.

The vehicle's environment currently consists of some laboratory open space in which a number of three-dimensional objects (some movable) can be placed in various configurations. To reduce requirements on the sensory and pattern-recognition system, the objects have simple geometric shapes.

There are enough different classes to permit specification of a variety of tasks. The objects may act as fixed or movable barriers and may form corridors or mazes. On the other hand, they may serve as the goals to be sought, examined, manipulated, or transported. There may be included two-dimensional signs or marks as part of the environment; these serve as landmarks for navigational purposes or for identification of features not easily represented otherwise.

We have purposely kept the vehicle and its environment as simple as we thought reasonable, in order to focus our attention on the conceptual and organizational problems that face the designer of an intelligent robot. These problems fall into the three major categories of perception (visual input and pattern recognition), natural-language understanding for command and control, and cognition (model building, planning, and monitoring of performance). Each of these areas will be reviewed briefly.

Perception

In the area of perception, the problems center about the description of scenes that the robot's TV camera views. Describing a complex visual scene is a far cry from classifying a single object, such as an isolated printed character in an optical character reader. In what language should the description of the scene be formulated? How does the computer recognize objects that are partially hidden from view? How does it describe objects that are only partially or ambiguously recognized? How does it perceive and define relations among objects, such as "the desk is near the door"? Each of these is a problem of mechanized perception and, equally, of software organization.

We have devoted considerable effort to the treatment of these perceptual questions.[5-7] Our current vision software is an amalgam of edge-enhancement techniques, heuristic routines for finding lines and regions, fixed and probabilistic decision trees, and context-aided scene analysis. Our programs are able to find simple objects and room boundaries in the scenes viewed by the vidicon camera and, with the use of projective geometry, to contribute to the robot's model of the floor plan of its environment. We have explored the addition of color, stereo, and texture analysis to the vision capability. Color appears most promising for the enhancement of recognition ability, and we plan to add color input in the future. A paper by Duda and Hart in these proceedings describes some of our recent work in visual scene analysis.[7]

The robot's present visual abilities are admittedly primitive in comparison with the rich visual world in which we live. While this is due in large measure to the inherent difficulty of visual scene analysis, it also derives in part from the crudity of existing video input equipment relative to the wondrous performance of the human eye. To take one example, the illumination levels in an office environment vary by more than 10,000 to 1, and the human eye encompasses this range and more with ease. In the future, the evolution of better input hardware and better visual-recognition software can be expected to proceed together and reinforce each other.

Natural-Language Understanding

In the area of command and control, the robot must be able internally to make "sense" of commands from a human operator. When humans communicate with each other, some information is redundant; other information is missing, and is supplied by the listener. A successful robot must allow for these factors. If you had a robot as a household assistant and told it to "make the bed," you would want it to interpret this command as applying only to beds that were not already made up. Also, you would want it to have the sense not to head for the garage for tools and lumber to start building a bed! Ultimately, the reservoir of implicit knowledge that the robot is asked to bring to bear in performing its tasks may compare with that which a human acquires in his childhood--a staggering requirement in terms of today's computer capabilities.

In the present SRI robot system, the operator or experimenter communicates with the robot through a teletype unit attached to the computer. Rather than force the operator to use the internal computer language of the robot program, we have taken a first step in the direction of natural-language recognition outlined above.[8,9] The operator enters statements and instructions for the robot in a simplified subset of English, and a natural-language interpretation routine converts these statements and instructions into the mathematical form (predicate calculus) in which they are represented in the computer. The capability of this interpreter is steadily augmented through the laborious process of adding software routines for the recognition of new word types, new sentence structures, and so on. This natural-language interpreter could easily be adapted to handle spoken input, given an acceptable state-of-the-art in speech recognition.

Cognitive Problems

Many of the research problems fall into the overlapping areas of model building, planning (or problem solving), and monitoring. The cognitive system of the robot must build an internal model of the state of the vehicle and its environment, which it can access to determine the external state of affairs. Using this model, the system must be able to make plans--beginning with the goal set by the operator, to develop a structure of subgoals, sub-subgoals, etc., down to the individual actions that it will attempt to execute. Finally, the robot must be able to monitor the attempted execution of its plans--to keep the internal model in agreement with the external situation and to discover and react when the progress of the execution diverges from the plan. In a complex real environment, which is too rich to be modeled precisely, such divergences are bound to occur; "the only thing certain is uncertainty." Planning, monitoring the execution of the plans, and making revisions necessary to the completion of the assigned tasks in a real environment, together constitute probably the ultimate challenge in robotry research.

Most of our work to date in this area has centered on the use of a question-answering program, called QA3.[10-13] A goal to be achieved by the robot is transformed (by the natural-language interpreter) into a question to be answered: "Does there exist a state in which such-and-such a condition is true?" This question is cast as a theorem in predicate calculus to be proved by the robot, using "axioms" that describe the operations the robot can perform and their consequences. This method is very powerful because the chain of axioms that the robot uses to prove a theorem specifies at the same time a plan for the robot to carry out in attempting to achieve the goal. The work on question answering in our group has exhibited an exciting interaction between three areas of current interest in artificial intelligence: robotry, information retrieval, and theorem proving by machine.

While QA3 has allowed us to use powerful formalized methods of artificial intelligence for the planning function of the robot, it is weak with regard to monitoring the execution of the plans and coping with the uncertain and the unexpected. We are currently exploring several avenues for the development of a more capable robot.[14,15] One of these is a more powerful theorem prover; other approaches involve various probabilistic, goal-seeking, and feedback-controlled structures for the cognitive monitor system of the robot.

Because the problems of perception, control, and monitoring in the real world are so formidable and so little explored, the day when we entrust a robot with operation in complex and unregulated environments (such as driving, lumbering, or building construction) lies at the end of a long trail of both basic and applied research. We expect to find applications of our work sooner in surroundings that can be largely controlled (warehouses, libraries, assembly lines, machine shops) and in those where it is unpleasant, difficult, or dangerous for a man to operate directly. The mode of operation of a robot can vary from step-by-step human guidance to complete autonomy. In any case, our work is devoted to ultimately enhancing the robot's abilities by allowing it to perceive, plan, and act independently and automatically with regard to some portion of its environment.

REFERENCES

1. Nils J. Nilsson, "A Mobile Automaton: An Application of Artificial Intelligence Techniques," Proc. International Joint Conference on Artificial Intelligence, Washington, D.C., May 7-9, 1969, pp. 509-520.

2. N. J. Nilsson and B. Raphael, "Preliminary Design of an Intelligent Robot," Computer and Information Sciences-II (Academic Press, Inc., New York, 1967).

3. Charles A. Rosen and Nils J. Nilsson, "An Intelligent Automaton," IEEE International Convention Record (1967).

4. C. A. Rosen, "Machines That Act Intelligently," Science Journal, pp. 108-114 (October 1968).

5. George E. Forsen, "Processing Visual Data with an Automaton Eye," Pictorial Pattern Recognition, Proc. Symposium on Automatic Photo-interpretation, pp. 471-502 (Thompson Book Company, Washington, D.C., 1968).

6. Claude R. Brice and Claude L. Fennema, "Scene Analysis of Pictures Using Regions," a paper submitted for publication in the Artificial Intelligence Journal; Technical Note 17, Artificial Intelligence Group, Stanford Research Institute, Menlo Park, California (November 1969).

7. Richard O. Duda and Peter E. Hart, "Experiments in Scene Analysis," Proc. First National Symposium on Industrial Robots, Chicago, Illinois, April 2-3, 1970.

8. L. S. Coles, "Syntax Directed Interpretation of Natural Language," Carnegie-Mellon University, Pittsburgh, Pa., Ph.D. Thesis (June 1967).

9. L. Stephen Coles, "Talking with a Robot in English," Proc. International Joint Conference on Artificial Intelligence, Washington, D.C., May 7-9, 1969, pp. 587-596.

10. Cordell Green, "Application of Theorem Proving to Problem Solving," Proc. International Joint Conference on Artificial Intelligence, Washington, D.C., May 7-9, 1969, pp. 219-240.

11. Cordell Green, "The Application of Theorem Proving to Question-Answering Systems," Stanford University, Stanford, California, Ph.D. Thesis (June 1969).

12. Bertram Raphael, "Programming a Robot," Proc. IFIP Congress 68 (North Holland Press, to be published).

13. C. Cordell Green and Bertram Raphael, "The Use of Theorem Proving Techniques in Question-Answering Systems," Proc. 23rd National Conference, ACM, pp. 169-181 (Brandon/Systems Press, Inc., Princeton, N.J., 1968).

14. R. E. Fikes, "A Heuristic Program for Solving Problems Stated as Nondeterministic Procedures," Carnegie-Mellon University, Pittsburgh, Pennsylvania (November 1968).

15. Peter E. Hart, et al., "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," IEEE Trans. on System Science and Cybernetics, Vol. SSC-4, No. 2, pp. 100-107 (July 1968).

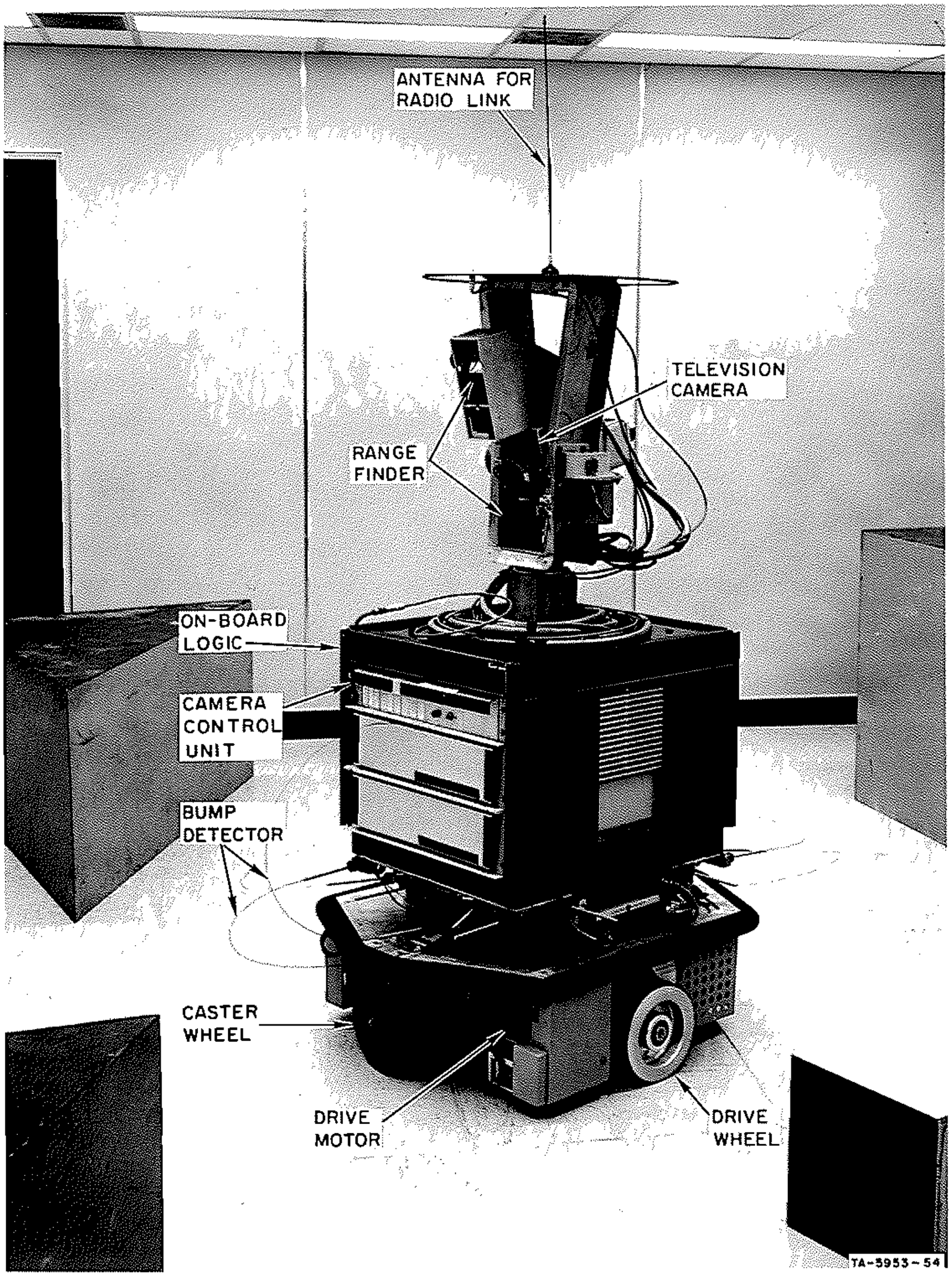

Fig. 1 The SRI Robot Vehicle