How Do Robots Take Two Parts Apart?

Ruzena K. Bajcsy and Constantine J. Tsikos

Department of Computer and Information Science
University of Pennsylvania
200 South 33rd Street
Philadelphia, PA 19104-6389

January 31, 1989

Abstract

This research is a natural progression of our efforts which begun with the introduction of a new research paradigm in Machine Perception, called Active Perception. There we have stated that Active Perception is a problem of intelligent control strategies applied to data acquisition processes which will depend on the current state of the data interpretation, including recognition. In this paper we treat the disassembly/ assembly problem as an Active Perception problem, and we present a method for autonomous disassembly based on this framework.

1 Introduction

Perceptual activity is exploratory, probing, searching [1], [2]. Percepts do not simply fall onto sensors as rain falls onto the ground. We do not just see, we look. And in the course of looking, our pupils adjust to the level of illumination, our eyes bring the world into sharp focus, our eyes converge or diverge, we move our heads or change our position to get a better view of something, and sometimes we even put on spectacles.

For robotic systems, this Active Perception approach has several consequences:

1. If one allows more than one measurement to be taken, then one must consider how they should be combined. This is the multi-sensory integration problem.

2. If one accepts that perceptual activity is probing and searching, then data evaluation techniques must be used to measure how well the system is accomplishing its perceptual task and to determine whether a feedback mechanism is needed.

3. If one accepts that perceptual activity is exploratory, then one must determine what must be built into the system in order to perform the exploration, i.e., what is a priori and what is data driven?

The next development in our program was the realization that perception is not only sensing but also involves manipulation [4]. For example, consider the problem of a static scene segmentation. This has been shown convincingly in our recent work [13] and in the paper: "Segmentation via Manipulation" [14] where we argued that a static scene that contains more than one object/part most of the time cannot be segmented only by vision or in general by any non contact sensing. Exception to this is only the case when the objects/parts are physically separated so that the noncontact sensor can measure this separation or one knows a great deal of a priori knowledge about the objects (their geometry, material, etc.). We assume no such knowledge is available. Instead, we assume that the scene is reachable with a manipulator. Hence the problem represents a class of problems of segmentation that occur in an assembly line, bin picking, organizing a desk top and their like. The typical properties of this class of problems are:

1. The objects are rigid. Their size and weight is such that they are manipulable with a suitable end effector. The number of objects in the scene is such that each piece can be examined and manipulated in a reasonable amount of time, i.e. the complexity of the scene is bounded.

2. The scene is accessible to the sensors, i.e. the whole scene is visible, although some parts may be occluded, and reachable by the manipulator.

3. There is a well defined goal which is detectable by the available sensors. Specifically the goal maybe: an empty scene, or an organized/ ordered scene.

The segmentation problem as is specified above is a sub-class of the more general disassembly problem, i.e. taking things apart which may be viewed as a process of getting insight into how to assemble objects, i.e. how to put pieces together. It is not difficult to see that this is how children learn about part/whole relationships and in general about an assembly process. But the question still remains; what perceptual information should be stored when such disassembly process takes place and is it enough for performing the assembly, i.e. the reverse tasks? This problem is what we call the Machine Perceptual Development and is at the heart of this paper.

One may ask how is Machine Perceptual Development related to machine learning? Relevant work on machine learning can be divided into two categories. One involves the application of the neural network paradigm, the other is studies of learning in the AI tradition. The neural net paradigm addresses problems at the low-level perception, learning patterns from the signal, but this approach does not answer the questions of data reduction from a signal that we are proposing. Moreover, we are trying to determine a useful division between "innate" structure and learned properties, that is to say, between a priori and data driven information. The traditional AI approach to learning has most frequently relied too much on

a priori information and has neglected the data driven part. We believe that this approach is too limiting.

2   The Two Part Disassembly Problem

We begin with the problem of the two part disassembly. The overall flow diagram of our methodology is as follows: Calibration/Exploration, Disassembly, Assembly. The fundamental issue is the REPRESENTATION. The case still has to be made for new representations that develop during an activity and that respect both the sensory apparatus and the task. **Traditionally, the Computer Vision community** has experimented with geometric CAD models for analysis, arguing that if CAD models are useful for making objects, then they should be equally useful for recognizing them. But such an argument is questionable. A designer creates a CAD model by specifying surface representations with detailed boundaries and explicit dimensions. To represent the internal dimensions, s/he shows cross sections. Finally, s/he specifies both the material and finish of the surface. Thus CAD models reflect how to synthesize an object during both its design phase and its manufacture.

The question is whether this same representation is useful for robotic analysis, i.e., object recognition necessary for disassembly and assembly. We believe the answer is no. First, the limits of sensors determine the limits to which a robotic system can differentiate between different materials, different colors, etc. A robot may not even have the sensors necessary to measure some of the properties that the designer has specified. For example, to distinguish metallic and non-metallic materials, a sensor is needed to measure conductivity. Secondly, the spatial resolution of a sensor limits how well a robotic system can measure spatial details: there is no point in representing a dimension of curvature with tight tolerances if a sensor cannot discriminate it. Thirdly, the noise of the perceptual system determines the minimal discriminability between different categories of objects. Finally, the robot may not know the substance/material of the object it is sensing. Hence it must have an apparatus to find such things out.

What follows in the subsequent sections is: First, the description of the Calibration process which will determine the physical and some geometric characteristics of the material (hardness, coefficient of friction, surface texture, conductivity , spectral properties such as reflectivity, weight/ density and their like). Second, the description of disassembly process and the division of build in procedures versus data driven part. Finally, the test of memory via assembly process.

2.1   Calibration/Exploratory Procedures

Unlike much of the current robotics effort we do not assume a priori knowledge of the physical nor geometric properties of objects that we deal with. In order to find out one must have build in capabilities, called Exploratory Procedures (EPs) [9] that seek out different

physical attributes. For this work we shall consider the following EPs: EP that determines the surface reflectance, discriminates between lambertian and highly reflective surfaces [3], EP for determining the hardness of the material and surface texture [12]. Notice that these EPs are static tests, i.e. the object is not manipulated. These EPs will give us the expected range of values for hardness, surface reflectance and surface texture. In the future we will add more attributes, such as electrical and thermal conductivity, measure of elasticity and deformability [11]. Furthermore, weight and density of the material as well moving parts, like objects on hinges, will be explored in a dynamic mode.

2.2 The Disassembly/Assembly System

First we shall describe the hardware configuration also shown in Figure 1. For the disassembly/assembly task, the robot is a six degree freedom PUMA 560 manipulator equipped with a range finder and/or a pair of CCD cameras, called the LOOKER and another six degree freedom PUMA 560 manipulator and a hand, called the FEELER. The LOOKER, depending on the need, can also have a color camera system or other non-contact electromagnetic wave measuring detector (infrared is one possibility). The FEELER has a force/torque sensor in its wrist and hand. The hand has three fingers and a rigid palm. Each finger has one and a half degrees of freedom. The sensors on the hand are: Position encoders, force sensors at each joint of the finger, tactile array at each of the finger tip and on the palm, Thermal conductivity sensor on the palm, ultrasound sensor on the outside of the hand. In addition, the hand has access to various tools that it can pick up under its control. Both of the FEELER and the LOOKER are under software control of strategies for data acquisition and manipulation. What are the Logical Components of the System? They are:

1. SENSOR MODELS that describe: The range of admissible values, the noise which determines the resolution, the geometry which determines the accessibility of the sensor to the investigated object or of its part.

2. TASK MODEL: In this case: a two part decomposition/separation.

3. PARAMETERS: About the physics/geometry of an object obtained through calibration EPs.

4. MANIPULATION PROCEDURES: such as: Push, Pull, Lift, Press, Turn, Twist, Grasp, Squeeze.

5. GEOMETRIC PROCEDURES: Shape description, especially detection of discontinuities, where is the binding force, size (length, area, volume) determination.

6. CONTROL STRUCTURE: (State, Actions), Priorities if more than one possible action, (here one may consider some cost/benefit function to make the right choice). Priority of sensing: how to start? (here we start with vision!). Detection of the goal state, i.e. two separate parts.

The Block diagram reflecting the logical components for disassembly/assembly is shown in Figure 2. This diagram is very similar to the one used in Tsikos's Ph.D thesis [13] for segmenting a complex scene. We have shown that:

1. Segmentation of an arbitrary scene requires not only a visual sensor, but also some manipulation actions, such as pushing, pulling, grasping and their like.

2. The interaction between the sensors and manipulation and the scene can generally be sufficiently modeled by a finite state, non-deterministic Turing Machine.

3. The critical consideration is the testability of the goal state. (In Tsikos' case it was an empty scene.)

2.3   The Disassembly Process

As a test for our system, consider a peg-and-hole problem shown in Figure 3. It is a test bed with the same shapes of the top of the peg but with differing holes (square, circular, or none) Figure 3d, 3c, 3a and with varying surface finish of the peg (smooth as shown in Fig. 3c and 3d, and threaded as shown in Figure 3b. This fixture has been designed so that we can test several combinations of manipulative actions. The general priority schema of control is as follows:

1. LOOK. Remember: Position and shape. Start with vision, identify the surface discontinuity of the peg-head vis-a-vis the hole surface, find the position, orientation, surface normal, and shape of the peg-head.

2. GRASP. Remember: Position and grasping force. After vision follow up with grasping in preparation for manipulation. The grasping procedure includes the limitations of the end-effector, i.e. this procedure utilizes the parameters obtained through calibration EPs and from the previous step which provides information on geometry of the peg-head.

   Our initial experiments were carried out using a parallel jaws gripper instrumented with force/torque sensors and tactile arrays. The goal of the grasp action is to verify correct grasp of an orthogonal parallelepiped peg-head. We define correct grasp to be a two PLANE contact between the jaws and the peg-head such that the forces and torques exerted on the sensors are of approximately equal magnitude and opposite sign.

   We use a binary search procedure in three space to verify and/or correct the position, orientation, and surface normal as computed by vision. The first step in this procedure is to make an initial grasp of the peg-head. In the general case the initial grasp will be two POINT contacts between the gripper jaws and the surfaces of the peg-head, See Figure 4a. Then we measure forces and torques. Using the sign and magnitude of these mesurements we un-grasp the object, reorient the gripper and attempt another grasp

until we have (in general) a two LINE contact between the jaws and the peg-head, as illustrated in Figure 4b. At each iteration, the changes in gripper orientation are one half of the previous step. This procedure continues until we have a two PLANE contact, see Figure 4c, and the forces/torques are of equal magnitude and opposite sign.

3. MANIPULATE-PULL. Remember: Direction and magnitute of pulling force while in the hole, and the positions during the departing motion (change in the magnitude of the pulling force). This procedure adaptively (using force feedback) pulls the peg by finding the direction which minimizes the reactive force. This procedure uses differential force feedback to subtract the grasping forces, recorded during the grasping phase.

4. OBSERVE the action using vision during manipulation. Remember: Shape, size and position of the two separating parts. An alternative to using vision during manipulation is to use a move until free primitive action that moves the manipulator slightly in a direction normal to the pulling force. If the disassembly of the peg is not yet complete then forces/torques will be exerted on the sensors. The system then returns back to the previous state, and continues with the manipulate- pull action.

5. GOAL STATE CHECK. If the two parts are separated then the goal state has been reached and stop. Notice that there are two ways to measure that the goal state has been reached. One is to use information from the contact sensors i.e. move until free, and the other is to use vision during manipulation to detect separation. In this work we use the former and we plan to integrate the latter soon. Notice that both methods allow us to measure the unknown length of the peg. Only vision, however, can measure the shape of the peg as well as the shape of the hole after the disassembly is complete. This is important in the general disassembly problem.

2.4 The Assembly Process

The fundamental question in disassembly is: Did the system remembered enough? Consider reversing the above described process: The FEELER is holding the head of the peg and we have stored the position and shape of the hole. Hence unless something has changed the FEELER can approach the hole without the LOOKER. The insertion process is the reverse of manipulate-pull. The goal state is determined by the length of the peg, that was remembered by the LOOKER after separation of the two parts. We conclude that at least in this test case the system remembered enough to pass the test.

3 Conclusion

We have defined and outlined our long-term thinking and investigations on Machine Perception that leads us to the latest research program of understanding (Machine Perceptual

Development). This is an outgrowth of our research on Active Perception, which views perceptual activity as an active process of SEEKING INFORMATION. Naturally this is not just blind pickup of any information. The system must protect itself by imposing some economy rules [8]. Even if the perceptual system receives overabundant amounts of information, again for economy reasons it must be selective in what it stores. Hence the fundamental problem remains: The REPRESENTATION issue. What is it that the system must have to seek, measure, and select in order to be able to move and manipulate?

Somewhat similar ideas appear in the work of Donald [5-7], and Pertin-Trocaz and Puget [10]. They consider a manipulation program automatically generated by a planner according to spatial and geometric criteria and ignoring uncertainities. Such a program is correct only if, at each step, uncertainities are smaller than the tolerance imposed by the assembly task. They propose an approach which consists in verifying the correctness of the program with respect to uncertainities in position and possibly modifying it by adding operations in order to reduce uncertainities. These two steps based on a forward and a backward propagation borrowed from formal program proving techniques are described in a general framework suitable for robotic environments. Forward propagation consists in computing successive states of the robot world from the initial state and in checking for the satisfaction of constraints. If a constraint is not satisfied, backward propagation infers new constraints on previous states. These new constraints are used for patching the program.

However, we differ in more than one ways from their approach. The most important difference is the ultimate goal, that is we are interested in the perceptual data reduction mechanisms rather than in a general plan of a process. We have posed these questions in the framework of disassembly of one object into two parts and tested the selected, remembered representation by reversing the process, i.e assembly. Our results are only very modest but we believe that they are encouraging!

4 Acknowledgements

5 References

1. Bajscy, R. K. (1982) What Can we learn from One-Finger Experiments. U.S. - France Seminar in Robotics, Paris, May, 1982.

2. Bajscy, R. K. (1988) Active Perception. Proceedings of the IEEE on Computer Vision, August, 1988.

3. Bajscy, R. Wohn, K. and Lee, S. W. (1988) Exploratory Procedures For Computer Vision. Submitted to the IEEE Transactions on Systems, Man and Cybernetics. October, 1988.

4. Bajcsy, R. K. and Tsikos C. J. (1987) Perception Via Manipulation. R. Bolles (Editor), 4th ISRR, Santa Cruz, California, August 9-14, 1987.

5. Donald, B. R. (1987). Towards Task-Level Robot Programming. Techical Report 87-878, Computer Science Dept., Cornell University, 1987.

6. Donald, B. R. (1988) The Complexity of Planar Compliant Motion Planning under Uncertainty. Proc. ACM Symposium on Computational Geometry, Urban, Ill., 1988.

7. Donald, B. R. (1988). Planning Multi-Step Error Detection and Recovery Strategies. Proc. IEEE International Conference on Robotics and Automation, Philadelphia, PA., April, 1988.

8. Hager G. (1988) Active Reduction of Uncertainty in Multi-Sensor Systems. Ph.D. Dissertation, Computer and Information Science Department, University of Pennsylvania, July, 1988.

9. Klatzky R. L., and Lederman S. There is more to touch than meets the eye: The Salience of Object Attributes for Haptics with and without Vision. Journal of Experimental Psychology, Vol. 116, No. 4, 1987. pp. 359-369.

10. Pertin-Troccaz, J. and Puget, P. (1987). Dealing with Uncertainity in Robot Planning using program proving techniques. R. Bolles (Editor), 4th ISRR, Santa Cruz, CA, August 9-14, 1987.

11. Sinha P. (1989) Haptic Exploration for Robots. Personal Communications, January, 1989.

12. Stansfield, S. (1987). Visually-Guided Haptic Object Recognition. Ph.D. Dissertation, Computer and Information Science Department, University of Pennsylvania, October, 1987.

13. Tsikos, C. J. (1987). Segmentation of 3D Scenes Using Multi-Modal Interaction Between Machine Vision and Programmable Mechanical Scene Manipulation. Ph.D. Dissertation, Computer and Information Science Department, University of Pennsylvania, December, 1987.

14. Tsikos, C. J. and Bajcsy, R. K. (1988). Segmentation Via Manipulation. Submitted to the IEEE Journal of Robotics and Automation, June, 1988.

"LOOKER"

"FEELER"

Laser
Range
Imager
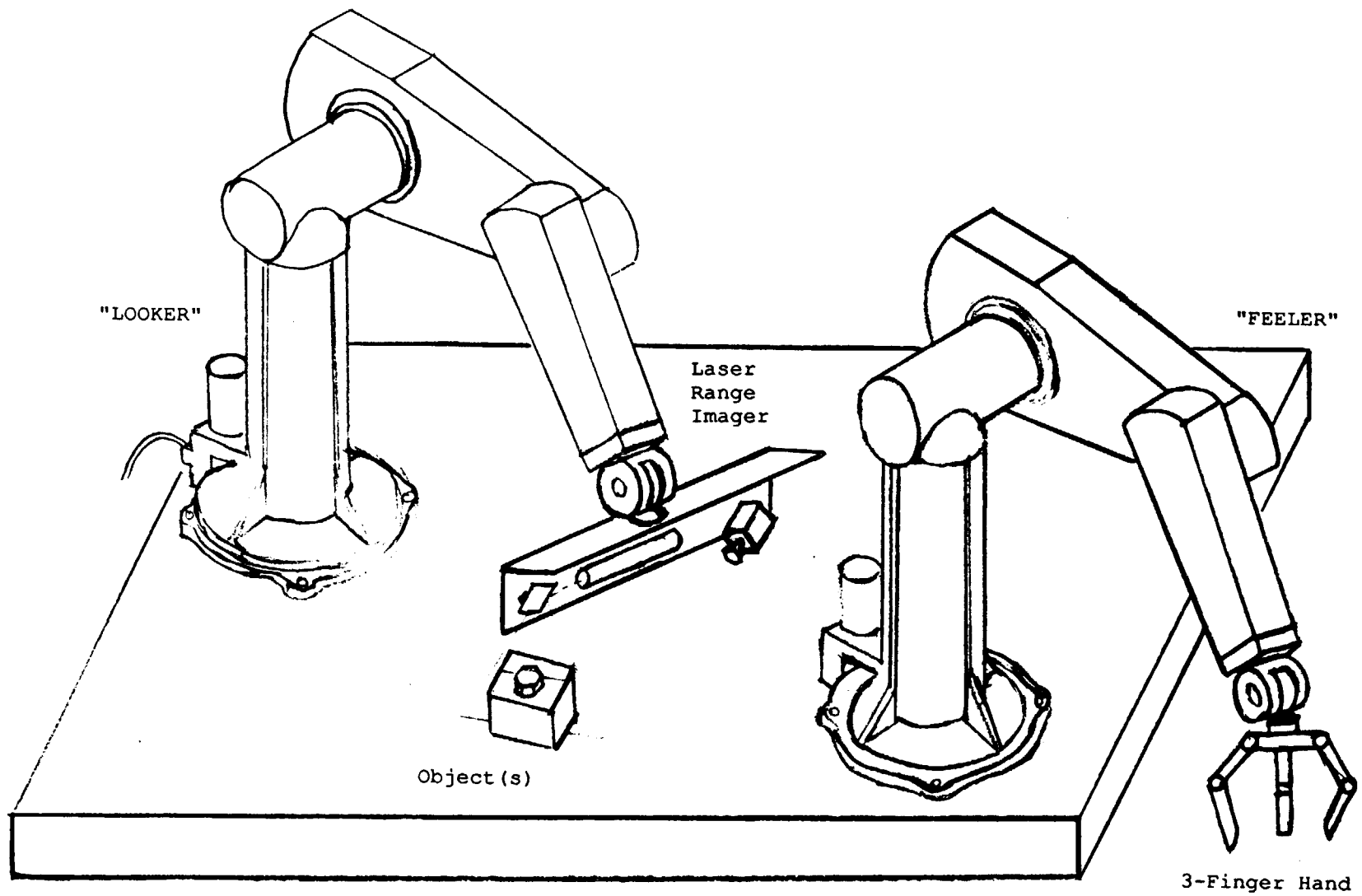
Object(s)

3-Finger Hand

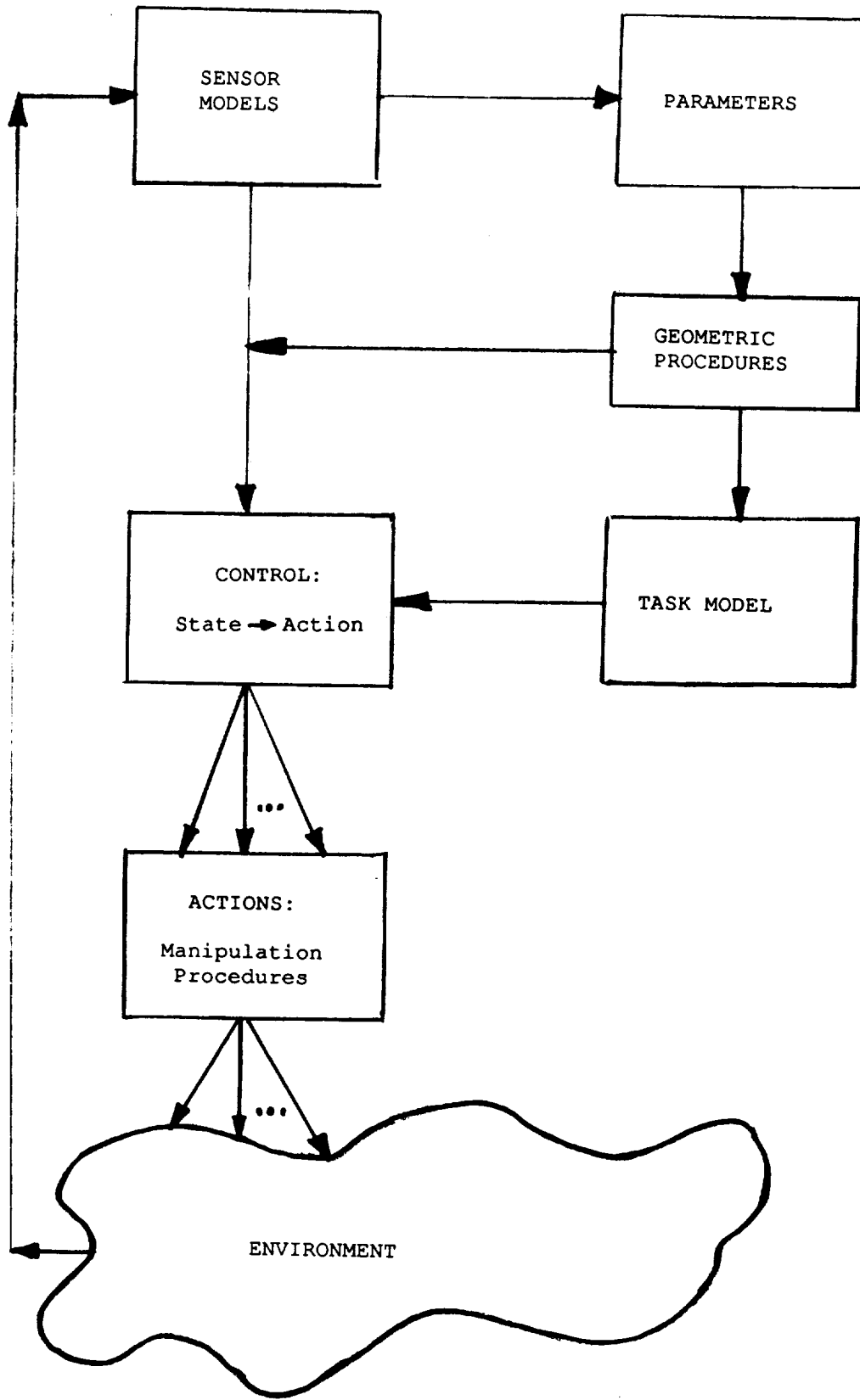Figure 1.   The Disassembly/Assembly System Hardware.

Figure 2. The Disassembly/Assembly System Block Diagram.
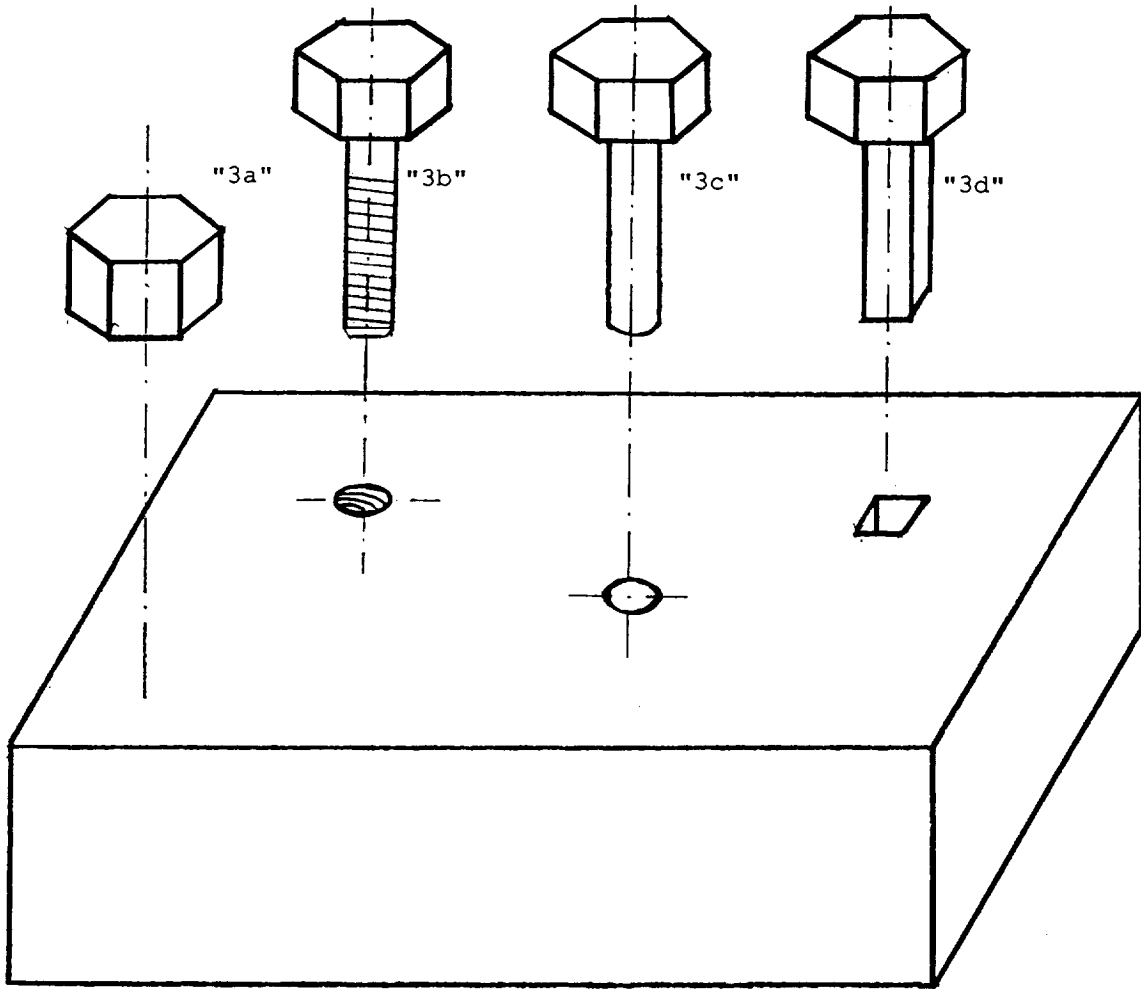
Figure 3. Various Instances of the Two-Part Disassembly/Assembly Problem.
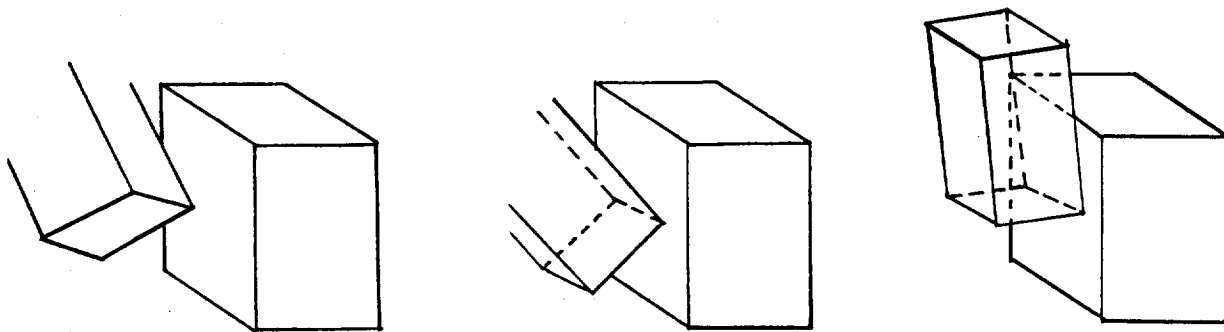

Fig. 4a. POINT Contact.   Fig. 4b. LINE Contact.   Fig. 4c. PLANE Contact.

Figure 4. Grasp with POINT, LINE, and PLANE Contacts. (Only one finger shown).