University of Pennsylvania

## ScholarlyCommons

Technical Reports (CIS)

Department of Computer & Information Science

January 1989

# Assembly via Disassembly: A Case in Machine Perpetual Development

Ruzena K. Bajcsy
*University of Pennsylvania*

Constantine J. Tsikos
*University of Pennsylvania*

Follow this and additional works at: https://repository.upenn.edu/cis_reports

# Assembly via Disassembly: A Case in Machine Perpetual Development

## Abstract

This paper presents first results in our effort of learning about representations of objects - The questions we are trying to answer are: "What is innate and what must be derived from the environment." The problem is casted in the framework of disassembly of an object into two parts.

## Comments

# ASSEMBLY VIA DISASSEMBLY:
# A CASE IN MACHINE
# PERCEPTUAL DEVELOPMENT

*Ruzena K. Bajcsy and*
*Constantine J. Tsikos*

MS-CIS-89-04
GRASP LAB 173

**Department of Computer and Information Science**
**School of Engineering and Applied Science**
**University of Pennsylvania**
**Philadelphia, PA 19104**

**January 1989**

Assembly via Disassembly:
A case in Machine Perceptual Development.
by

Ruzena Bajcsy
and
Constantinos Tsikos

**Abstract.**

This paper presents first results in our effort of learning about representations of objects.The questions we are trying to answer are: "What is innate and what must be derived from the environment." The problem is casted in the framework of disassembly of an object into two parts.

## 1. Introduction.

This research is a natural progression of our efforts which began with the introduction of the new research paradigm in Machine Perception, called Active Perception [Bajcsy,1982,1988]. There we have stated that Active Perception is a problem of intelligent control strategies applied to data acquisition process which will depend on the current state of the data interpretation, including recognition.

Perceptual activity is exploratory, probing, searching; 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 *multisensory 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 involves also manipulation [Bajcsy,Tsikos 1987]. Consider the problem for example of a static scene segmentation. This has been shown convincingly in our recent work [Tsikos,1987] and in paper: Segmentation via Manipulation [Tsikos, Bajcsy 1988] 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 noncontact sensing. Exception to this is only the case when the objects/parts

1

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 on an assembly line, bin picking, organizing a desk top and their like. What are the typical properties of this class of problems?

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

2. The scene is accessible to the sensors, ie. the whole scene is visible, although some parts maybe excluded, 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 subclass of a more general problem of disassembly task ie. taking things apart which maybe viewed as a process of getting insight into how to assembly objects,ie. 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 assembly process. But the question 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 task? 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 grossly 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 wrong and too limiting.

In order to be systematic we begin with the problem of two part disassembly. (this may suggest binary partitioning possibilities). The overall flow diagram of our methodology is as follows [Bajcsy et al 1988]: 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 processing 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 stiff 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: 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 as reflectivity , weight/density and their like).

Description of disassembly process and the division of build in procedures versus data driven part and finally the test of memory via assembly process.

## 2. 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)[Klatzky,Lederman 1987] 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 [Bajcsy,Wohn,Lee 1988], Recently we have made progress in this area, using color information [Bajcsy,Wohn,Leonardis, 1989],that is we can separate the highlights and interreflections from the basic color. EP for determining the hardness of the material and surface texture [Stansfield, 1987]. 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 shall add more attributes, such as electric and thermo conductivity, measure of elasticity and deformability,[Sinha 1989]. Furthermore, weight and density of the material as well movable parts like objects on hinges must be explored in a dynamic mode.

## 3. Disassembly/Assembly System.

First we shall describe the hardware configuration also shown in Figure 1. For the disassembly/assembly task, the robot will be equipped by one six degree freedom manipulator and a range finder and/or a pair of CCD cameras, called the LOOKER and another six degree freedom manipulator and a hand, called FEELER. The LOOKER, depending on the need can also have a color camera system or any non-contact electromagnetic wave measuring detector(infrared as one possibility). The FEELER has a force/torques sensor in its wrist and a hand. The hand has three fingers and a rigid palm. Each finger has one and half degrees of freedom. The sensors on the hand

are: position encoders and force sensor at each joint of the finger, Tactile array at each of the finger tip and on the palm, Thermo-sensor on the palm, ultrasound sensor on the outer side of the hand. In addition the Hand has available various tools that it can pick up under its control.

Both of the FEELER and LOOKER are under software control of strategies for data acquisition and manipulation.

What are the Logical Components of the System? They are:

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

2. The Task Model:
   As was mentioned in the introduction, we limit ourselves to the task of two part decomposition/separation. We like to argue that this is a rather generic problem in many maintenance tasks, such as: remove and install, test and repair, and their like, see also [Siegel et al.,1982].

3. Parameters about the Physics/Geometry of the object obtained through calibration EPs.

4. Manipulatory Procedures:
   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 can 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, ie. 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[Tsikos,1987] for segmenting a complex scene. We have shown that

1. Segmentation of an arbitrary scene requires not only a visual sensor, but also some manipulatory 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 but non-deterministic Turing Machine.

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

As an test for our system, consider a peg-and -hole problem as 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 and none) Figure 3d, 3c, 3a and with varying surface finish of the peg (smooth as shown in Figure 3c and 3d, and threaded as shown in Figure 3b). This is so that we can test combinations of manipulative actions.

The general priority schema of control is as follows: START: VISION; REMEMBER: POSITION and SHAPE. Start with Vision, identify the surface discontinuity of the Peg-head vis-a-vie the hole surface, find the position of the peg-head and its shape. GRASP; REMEMBER: POSITION, GRASPING FORCE After Vision follow up with Grasping in preparation to manipulation. The grasping procedure, includes the limitations of the end-effector, that is this procedure utilizes the parameters obtained through calibration EPs and from the previous step which provides information on geometry of the peg-head; MANIPULATE-PULL; REMEMBER: DIRECTION, MAGNITUDE of PULLING FORCE while in the hole, POSITION of departing the hole (change in the magnitude of the pulling force); This procedure adaptively (using force feedback) pulls the peg by finding that direction which minimizes the reactive force; OBSERVE simultaneously using vision the manipulatory action. REMEMBER: SHAPE, SIZE and POSITION of the SEPARATED TWO PARTS.

GOAL STATE CHECK. If two parts separated then the Goal state has been reached and stop. If not then repeat the whole sequence again.

Assembly Process:
Did the System Remember 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 stick 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.

## 4. 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 some economy rules[Hager,1988]. Even if The perceptual system receives over abundant amounts of information, again for The 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 from The perceptual information in order to be able to move and

manipulate? Somewhat similar ideas appear in The work of Pertin-Trocaz and Puget [1987]. They consider a manipulation program automatically generated by a planner according to spatial and geometric criteria and ignoring uncertainties. Such a program is correct only if, at each step, uncertainties 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 uncertainties in position and possibly modifying it by adding operations in order to reduce uncertainties. 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 c 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 way from their approach. But 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 casted 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!

## References:

Bajcsy, R., Lee, S.W., & Leonardis, A. (1989).
Image Segmentation with Detection of Highlights and Inter-reflection Using Color. In Proceedings of The Image Understanding and Machine Vision Conference, North Falmouth,Cape Cod, MA. Optical Society of America, 1989 Technical Digest Series, Vol.14. pp.16-19.

Siegel, A.I., Barter, W.D., & Kopstein, F.F. (1982).
Job Analysis of Maintenance Mechanic Position for The Nuclear Power Plant Maintenance Personnel Reliability Model. Oak Ridge National Laboratory/TM-8301, Oak Ridge,Tennessee 37830

Pertin-Trocaz-Troccaz, J. & Puget, P. (1987).
Dealing with Uncertainty in Robot Planning using program proving techniques. In B.Bolles (Ed.) ISRR 87, Santa Cruz, California.

Donald, B.R. (1987).
Towards Task-Level Robot Programming. Tech.Report 87-878,Dept. of Computer Science, Cornell University, N.Y. 14853.

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

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

**Acknowledgment:**

"LOOKER"

"FEELER"

LASER
RANGE
IMAGER

OBJECT(S)
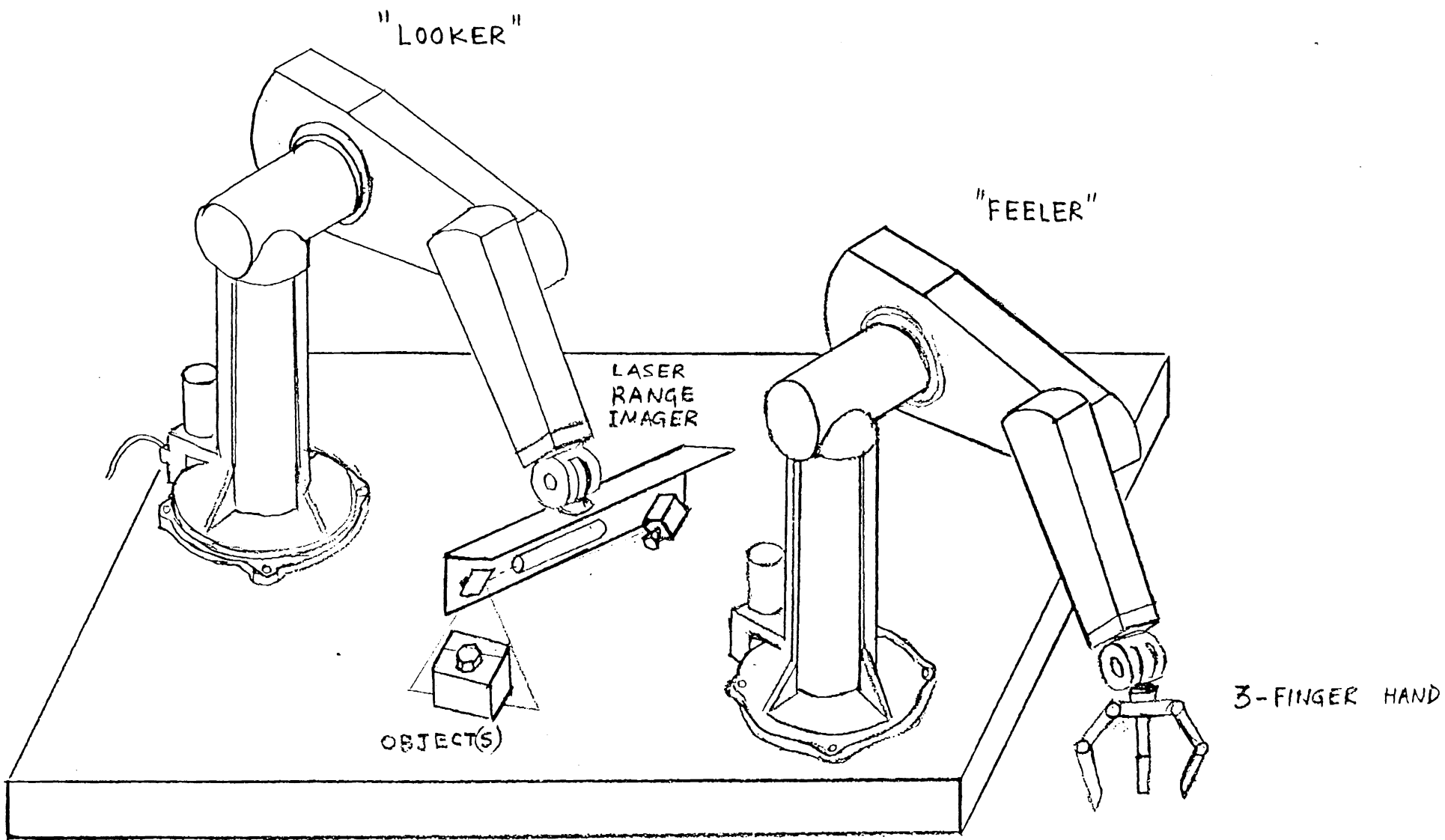
3-FINGER HAND

Fig. 1.    The Disassembly/Assembly    System   Hardware Implementation.
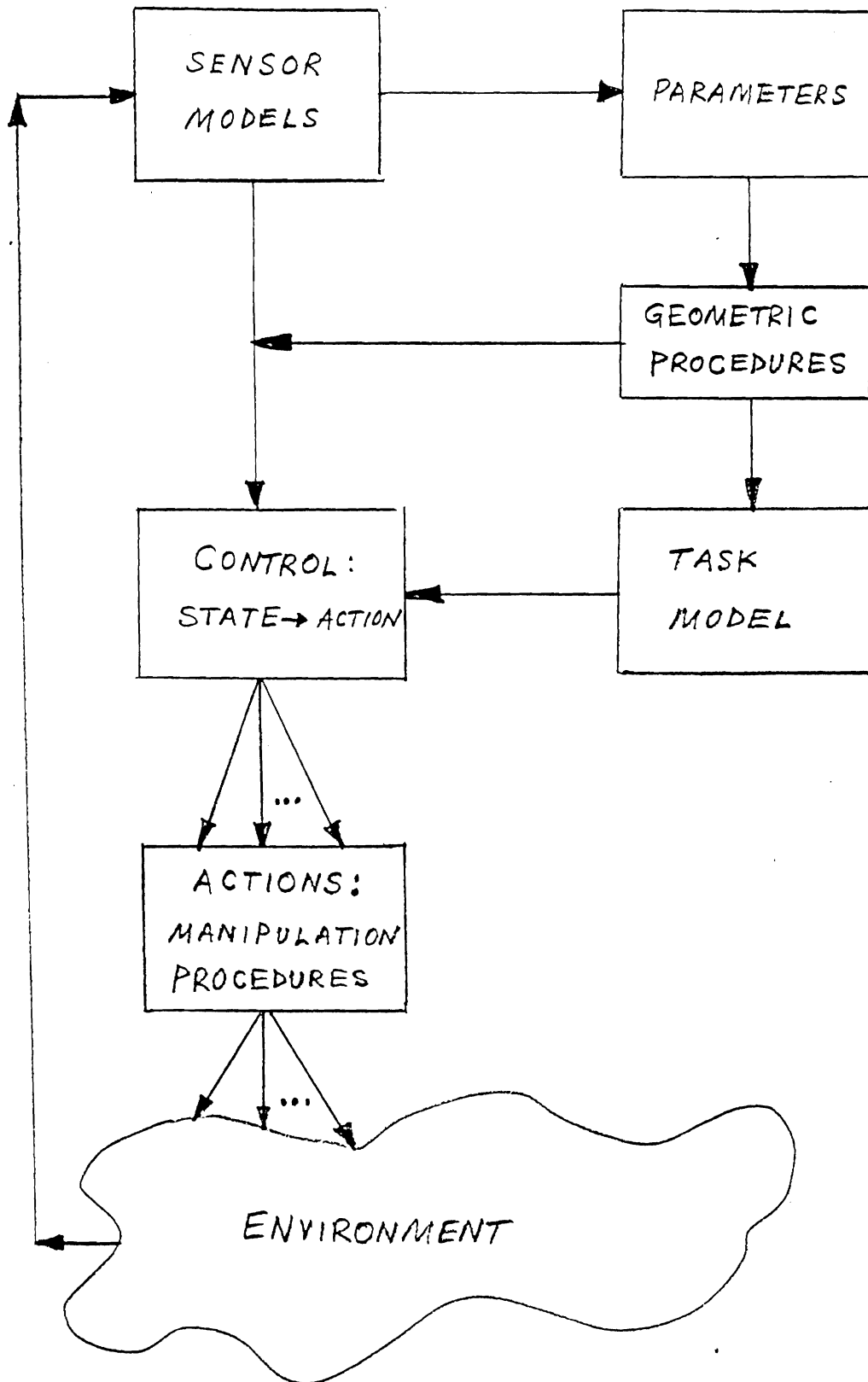
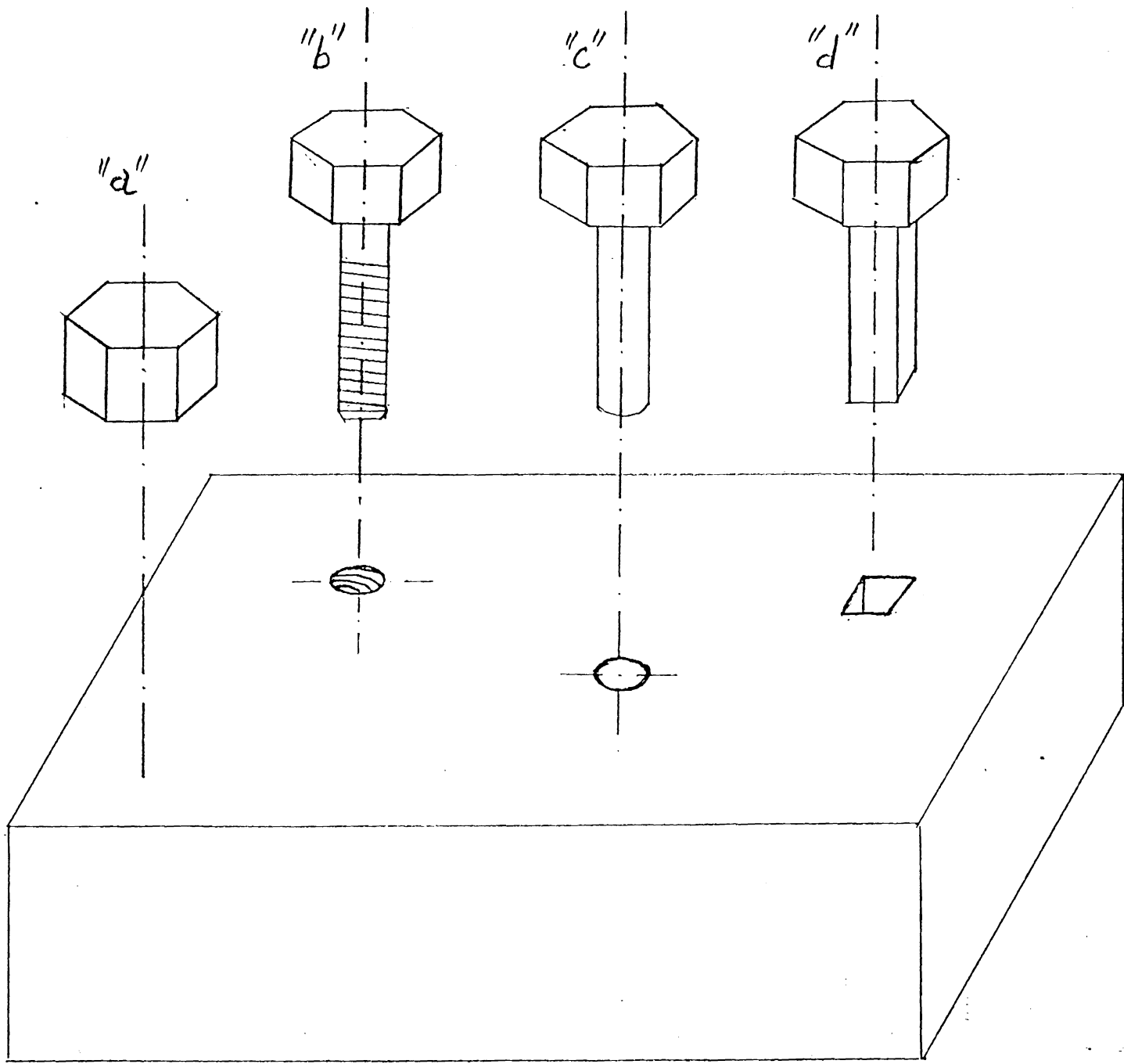Fig. 2    The Disassembly/Assembly System
Block Diagram.

Fig. 3. Various Instances of the Two-Part Disassembly/Assembly Problem.