SRI International

COMPUTATIONAL STRUCTURES FOR MACHINE PERCEPTION

Technical Note 233

January 1981

By: Martin A. Fischler
Program Director, Vision
Artificial Intelligence Center

This note is a summary of a talk given at the La Jolla Institute Workshop on Research Requirements for Advanced Computer Design (August 5-6, 1980). The purpose of this workshop was to identify important technical areas where radically new computer architectures might be necessary for future progress.

SRI International
333 Ravenswood Ave. • Menlo Park, CA 94025
(415) 326-6200 • TWX: 910-373-2046 • Telex: 334-486

ABSTRACT

This note discusses the adequacy of current computer architectures to serve as a base for building machine vision systems. Arguments are presented to show that perceptual problems cannot be completely formalized and dealt with in a closed abstract system. The conclusion is that the digital computer, organized as a general-purpose symbol processor, cannot serve as an adequate instrument for achieving a human-like visual capability.

CONTENTS

I INTRODUCTION

It is almost universally true that living organisms can interact intelligently with their surrounding physical environment, and that machines do not have this ability. This note will present arguments as to why our attempts to build machines with perceptual abilities have progressed so slowly, and indicate what developments in machine architectures will be needed to improve the existing situation.

II THE REPRESENTATION PROBLEM

The task of memorizing a list of 30 digits printed on a piece of paper, after a few seconds inspection, would probably be impossible for all but a very small number of people. On the other hand, an aerial picture of the Golden Gate Bridge could easily be memorized so that at some future time it could be distinguished from a variety of other scenes. Without quibbling about the exact nature of these two tasks, I think it would be agreed that the means by which we try to remember the information depicted in these two scenes is considerably different. In the case of the numerals, our memorization is primarily based on assigning a specific name (i.e., the names commonly assigned to the given numerals) to each depicted object. In the case of the natural scene, our memory is primarily that of a picture. It is apparent that at least two distinctly different types of representations are required for human modeling of the environment.

I will define "Symbolic Representation (SR)" to mean that the physical representation of the information is based on a purely arbitrary convention; for example, unless you knew that the binary number "111" is the encoding for the cardinal number seven, there is no way you could figure this out. I will define an "Isomorphic Representation (IR)" as one in which the physical representation of the information bears a resemblence (in either appearance, or function, or both) to the object or event being represented.

What does all this have to do with computer architecture? The conventional digital computer, as a general-purpose computing engine, has achieved its generality and power by imposing two demands on the problems it can deal with. First, that the data be represented symbolically, and second, that the transformations to be imposed on the data be defined in abstract algorithmic terms -- the physical nature of the devices that perform the computations are completely masked. For a wide variety of important problems this is exactly what we want, but for perceptual problems we cannot afford the price. Let us see why.

III THE INADEQUACY OF THE SYMBOLIC REPRESENTATION FOR PERCEPTUAL PROBLEMS

Since the SR is a completely arbitrary encoding of the information it represents, every relevent aspect of the situation being dealt with must be explicitly described; for example, even such "obvious" facts as: "if A is to the right of B, then B is to the left of A," must be explicitly entered as part of the data base for a computer program that must reason about spatial relations. For most practical problems, the volume of such explicit information is unreasonably large.

Not only is the above problem one of sheer volume of information, but there is also the requirement that one has a complete understanding, in symbolic terms, of all the relevant physical phenomena.

Finally, there is the gross inefficiency in processing speed that results when an inappropriate representation is used in a problem-solving context. For example, consider the problem of determining analytically (symbolically) whether a car could be pushed through a wooded patch of terrain, as opposed to actually trying it, or even simulating the task with a small-scale physical model.

IV THE INADEQUACY OF CONVENTIONAL COMPUTER ARCHITECTURES FOR SOLVING PERCEPTUAL PROBLEMS

The main thesis of this note, as presented above, is that perceptual problems cannot be completely formalized and dealt with in a closed abstract system. Thus, the digital computer, organized as a general-purpose symbol processor, cannot serve as an adequate instrument for perceptual reasoning.

In the remainder of this note I will address the additional issues of computation (complementary to, but distinct from those of representation), and after categorizing the computational requirements for perceptual reasoning, I will indicate what architectural innovations will be necessary to satisfy these requirements.

V CLASSIFICATION OF COMPUTATIONAL PROBLEMS FOR MACHINE PERCEPTION

* Formalization (hypothesis formulation, model definition, model selection): this is the act of recognizing that some segment of the environment has a coherent structure or purpose, and creating or assigning a symbolic description to represent the situation.

  -Example: recognizing that light is reflected from the surface of objects in different ways depending on the nature of the material composition of the surface (i.e., specular reflection, Lambertian reflection, etc.), and defining solvable mathematical models that can be used to predict the photometric appearance of objects, given their material type.

  -Difficulty: formalization is a creative act that includes a component of scientific discovery. General techniques for accomplishing this task are not currently available, nor do we understand this process well enough to even make a good guess as to how it might be accomplished.

* Transformation (execution of algorithmic procedures): given

a model (symbolic representation), we often wish to make explicit some implied aspect of the underlying object or event; this is accomplished by transforming the initial explicitly modeled information via some specified procedure.

-Example: convolving a high-pass filter with an image to enhance (make explicit) the visible edge structure.

-Difficulty: many of the transformations of interest do not have efficient algorithmic realizations on available computing machines; often the required processing time makes certain desirable transformations impractical to apply.

* Optimization (formal "internal" search): given a representation that defines (implicitly or explicitly) all possible alternatives for some universe of discourse, find the alternative that maximizes some objective.

-Example: find the shortest path between two specified points in a graph.

-Difficulty: models (and the associated objective functions) for many important problems have so many relevant parameters, that instantiation of all these parameters is impractical. Optimization procedures specified in terms of the operations available on conventional digital computers often have exponential complexity.

* Experimentation (informal "external" search; active or passive sensing of the environment): if we consider the external environment to be part of the data base (representation), then experimentation can be viewed as the process of accessing this information.

-Example: measure (rather than deduce) the light reflected from the surface of a given object.

-Difficulty: since we can only possess an incomplete internal representation of the external environment, the process of experimentation cannot be completely algorithmic, but must be adaptive, and possibly creative, to be effective.

* Interpretation: selecting and instantiating the available models that can be used to represent a given physical situation; resolving the conflicting overlapping assertions of these models.

  -Example: provide an English language description of a given scene.

  -Difficulty: incompatible, incommensurate, inaccurate, and incomplete models and information are often the rule in realistic perceptual tasks.

VI ARCHITECTURAL IMPLICATIONS FOR ACHIEVING MACHINE PERCEPTION

Almost all the problem areas discussed above could profit from machines with more memory and higher speed. In some of the cases, it is reasonable to expect that more effective extensions of current approaches will provide adequate solutions. However, if we wish to achieve human levels of perceptual performance, there are a number of areas where fundamental changes will probably be required in both the architecture of the computing device, and how it interacts with its external environment; some of these areas are discussed below.

* Multiple representations: it was argued earlier that the symbolic representation, by itself, was inadequate for dealing with many perceptual problems. The use of "isomorphic" representations and associated algorithmic techniques, probably by directly invoking the physical properties of specially selected computing hardware, will be required.

* Direct interaction with the environment (experimentation): real perceptual domains generally have the property that they are too complex to be adequately modeled, either because we do not fully understand the underlying physical principles and constraints, or because of the impracticality of explicit description even when understanding is present. The external environment must thus be directly accessible to the machine (via sensors and effectors) as part of its data base.

In addition to merely accessing the "external data base," there is also the critical issue of updating the system's model of its physical coupling to the external world, necessitated by dynamic changes occurring in both the physical structure of the system and in the characteristics of the surrounding environment; such updating must be based on a comparison of predicted versus actual results of interactive transactions. Sensory deprivation experiments provide an indication of the importance of this process for higher (living) organisms.

* Incomplete specification and learning: the role of learning in accommodating to unanticipated events is obvious. However, it may be the case that complete specification of a system capable of intelligent interaction with even an unchanging environment is impractical. One way of resolving this problem is to provide a partial specification of the system, and provide it with a capability to "mature" through direct environmental interaction. Nature has found it effective to take this approach, and intelligent machines may have to be designed the same way.

* Control structures: complex systems, interacting with complex (incompletely modeled) environments, cannot be designed to anticipate all contingencies. Further, they invariably will contain design errors, as well as being subject to mechanical failure. Conventional control structures for computing systems are inadequate in dealing with such problems. On the other hand, we observe that social, political, religious, military, and commercial entities have evolved effective control structures for organizing and allocating resources in an uncertain and changing environment. It appears reasonable to believe that computing systems, which can afford the luxury of redundant hardware and multiple algorithms with the same functional goals, can employ the principles of societal organizations to achieve reliable and intelligent behavior in the face of complexity and uncertainty.

VII CONCLUDING COMMENTS

The primary issue of representation for perceptual systems is not that of analog versus digital, but isomorphic versus symbolic; and the resolution of this issue will not depend on what is theoretically possible, but on the practicality of explicit specification of the immense data base needed for a purely symbolic approach, as opposed to capturing the needed knowledge implicitly via directly exploiting the physical characteristics of the computing machinery.

Machine architectures for perceptual tasks will require, in addition to new types of internal representations, more direct access to the external world, and control structures that can function in spite of hardware and design failure.