A. I. MEMO 366 **MAY 1976**

PROPOSAL TO THE

ADVANCED RESEARCH PROJECTS AGENCY

by

Patrick H. Winston

And The Staff Of

The Artificial Intelligence Laboratory

Of The

Massachusetts Institute Of Technology

1 January 1976 To 31 December 1976

Abstract

This is the substance of a proposal submitted in June, 1975, for research in the areas of large data bases and intelligent terminals, applications of machine vision and manipulation, basic studies in Artificial Intelligence, and LISP machine development.

Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research Contract number N00014-75-C-0643.

TABLE OF CONTENTS

THE WORK STATEMENT

THE WORK STATEMENT

This proposal consists of four major components:

- The Artificial Intelligence Laboratory proposes to work on issues central to the large data base problem, the intelligent terminal problem, and the natural language engineering problem with a broad, integrated research program. There are several imediate tasks:

  To produce a natural language interface for accessing intelligent terminal and large data base tools.

  To produce a natural language generating program to act with the understanding program.

  To develop Minsky's frame system into a form suitable for dealing with large bodies of knowledge.

  To create resourse allocation programs suitable for a variety of scheduling tasks.

  To create models of people and organizations such that programs can understand their needs.

- The Artificial Intelligence Laboratory proposes to work on applications of computer vision with the emphasis shifting from tool building toward solving real problems. Several new tasks are to be undertaken:

To create an inspection program for dealing with sophisticated jet engine castings.

To create representations for terrain such that vehicles can discover their own orientation using it.

To develop tracking programs that use knowledge about the objects tracked beyond simple velocity and acceleration prediction.

- The Artificial Intelligence Laboratory proposes to continue a basic studies program aimed at dealing with deep, central issues. The issues are to be exposed by work in computer vision, common sense reasoning, and expert problem solving. The ubiquitous representation problem is involved in all three areas. The basic goals include the following:

  To create solid symbolic representations from images.

  To create programs which can do common sense reasoning about the simple physical world.

  To create expert problem solvers that expose the general issues pervasive in the creation of all engineered systems.

- The Artificial Intelligence Laboratory proposes to complete development of low cost hardware intended to serve single users with an outstanding LISP computing system dramatically superior to any available today both in terms of power and cost.

There is to be a working prototype by June, 1976.

There is to be an advanced prototype by June, 1977.

INTRODUCTION TO THE PROPOSAL

THE PROPOSAL

This proposal specifies both our long term direction and our specific goals as planned for the next two years. We relate these goals to past progress, immediate practical needs, and long term scientific objectives. Since many of the scientific objectives cited are deep and demanding, readers should understand that considerably more than two years will be required for full solution.

THE FOUNDATIONS

From the beginning our purpose has been to make machines more intelligent, partly to make them more useful and partly to learn about the nature of intelligence. Certainly the need to make machines smarter is clear. It is also clear that the ideas of Artificial Intelligence are the principle scientific path toward making machines smarter. *How soon will it happen and how much will it cost?*

What makes prediction difficult, of course, is the non-linear relationship between the level of technical development and the payoff. For any combination of Artificial Intelligence paired with a need, there is some threshold on costs and capabilities such that on one side, little happens, and on the other, revolution. The pocket calculator phenomenon is one example of this that is known to everyone. The pocket computer is around the corner. We believe that there are many areas in which Artificial Intelligence will produce its own revolutions resembling that of the pocket calculator in their swiftness. Two particularly promising areas are these:

- The combination of large data base handling, intelligent interfacing of software tools and people, and natural language interface engineering.

This revolution will be based on a qualitative advance in which programs *understand*, to a significant degree, the information they are handling; the current technology merely applies manipulative rituals that do not use much of the knowldege embedded in the data bases being manipulated.

- The application of the ideas of artificial intelligence to computer vision and manipulation for productivity technology, vehicle guidance, and image understanding.

Again, the breakthrough will depend on the use of representations of the meanings, uses, and special features of the objects to be sensed and manipulated.

These are only what we view as immediate areas of maximum opportunity. Many others are suggested by the following sample:

SOME AREAS OF APPLICATION

Manufacturing	Assembling Electronic Systems
	Assembling Small Mechanical Devices
Mining	Coal Mining In Dangerous Mines
	Undersea Recovery Of Manganese Nodules
Farming	Selective Harvesting With Mixed Crops
	Pest Control
	Pruning
Repair And Maintenance	Vehicle Debugging
	Floor Care

Purposeful Monitoring	Intelligent Autopilots
	Shipboard Functions
Automatic Programming	Management Reports
	Inventory Control
	Production Scheduling
	Quality Control
Management Assistance	Very Large Data Bases
	Intelligent PERT Networks
	Scheduling People and Groups
	Question Referral
	News Summarizing
	Document Draft Polishing
Logistics	Routing
	Intelligent Substitution
	Equipment Sharing
Medicine	Diagnosis
	Intensive Care Monitoring
	Treatment Management
	Radiation Therapy Setup
Education	Intelligent CAI Systems
	Information Retrieval
Fire Fighting	Fearless Firemen
	Apparatus Allocation
Theft Prevention	Trackers

How can such objectives be reached? Certainly for smart machines to make an impact in areas like those mentioned, the smart machines must exhibit a variety of talents associated with intelligence. Computer vision, English understanding, learning and debugging, and expert problem solving all must be understood. It is natural then that considerable effort has gone into understanding these things. Happily we get back to them both by way of our desire to lay the foundation for applications and by way of our hope to understand intelligence for its own sake.

GENERAL OBJECTIVE: TO MAKE MACHINES SMARTER

TO MAKE MACHINES MORE USEFUL	TO UNDERSTAND INTELLIGENCE
Areas Of Application	Systems To Be Understood
Intelligent Terminals	Image And Scene Analysis
Very Large Data Bases	Speech And Text Understanding
Remote Guidance	Learning and Debugging
Photo Interpretation	Expert Problem Solving
Logistics	Common Sense Reasoning
Purposeful Monitoring	Goal Understanding
Repair And Maintenance	
Cleaning	
Manufacturing and Mining	
Medicine	
Education	
.	
.	
.	

As the diagram suggests, it has been sensible during the gestation years to be organized around themes like computer vision, English undestanding, learning and debugging, and expert problem solving. It is through these particular thrusts that we have learned a great deal about what are now regarded as central theoretical issues:

SOME CENTRAL THEORETICAL ISSUES

Knowledge Representation	Constraint Exploitation
Search And Control	Classification And Deduction
Programming Methodology	Goal Directed Versus Bottom Up
Grouping	Serial Versus Parallel

THE FUTURE

The Structure Of The Laboratory Is Evolving With The Science

The Artificial Intelligence Laboratory is widely recognized as one of the world's principal sources of ideas, experimental prototypes, and feasibility demonstrations on the frontier of computer science. It would not be appropriate for the laboratory to concentrate on producing operational systems in areas like that of very large data bases, but it does seem appropriate to begin thinking in terms of what AI can lead to in the near term and to begin producing the basic research results lying just in front of prototype production. To this end, we now see ourselves organized into a new structure that draws not only from the desire to understand various dimensions of intelligence but also from a more direct interest in applications challenges on one side and an interest in the emerging central issues on the other. Manifestly the boundaries are to be soft, not hard, and as has been common in the past, much flow among the particular subgroups is expected.

Applications Oriented Basic Studies

Because natural language research has made much progress along with parallel advances in our theories of knowledge, a considerable portion of our resources should be assigned to a unified program on very large data bases, intelligent terminals, and natural language interface engineering. Part of our efforts for the remainder of 1975 will be directed at bringing this area up to speed.

Applications Oriented Vision And Manipulation

We have spent considerable time and effort in developing a laboratory facility that is at once powerful, easy to copy, and relatively inexpensive. It has been developed using electronic circuit board assembly as the test problem. It is now time to redirect the use of this facility to other pressing problems by selecting the best of many alternatives for allocation of this important resource. Fully automatic assembly of small mechanical devices seems appropriate. Part of the image understanding problem also seems to be promising.

Issues Oriented Basic Studies

The study of basic issues with domains selected to facilitate understanding remains crucial to our program. Within this collection of topics we see continued work on the representation of knowledge, perhaps our most pressing problem area, as well as on expert problem solving, common sense physical reasoning, problem solving by analogy, learning, and constraint driven problem solving of the sort exposed by studies in computer vision. This last mentioned work on computer vision is believed to be critical for long range progress on visual interpretation problems that require nearly all of a large array of image points to be processed.

Systems And Systems Concepts

Progress in Artificial Intelligence is facilitated by adequate equipment. Indeed, many of its great accomplishments have involved programs that are too large and to complex to be produced without the frontier time sharing concepts we have developed. Now time sharing is well established, but newer, even more revolutionary opportunities have emerged. It therefore seems wise to continue our systems concepts leadership, both because of our own needs in Artificial Intelligence and because of the profound impact we believe these computer ideas will soon have on the computer consuming community.

RECENT PROGRESS

This introduction closes with a summary of recent progress organized around the topics which correspond to our past foci. These are:

Representing Knowledge and Expert Problem Solving
Understanding Natural Language
Understanding Vision
Productivity Technology
Learning and Debugging
Systems and Architecture
and New Programs

Sussman and his students have made progress in the direction of understanding debugging and the expert problem solving process in the particular domain of electronic circuitry. Sussman and Stallman, for example, completed a program and a paper in March of 1975, "Heuristic Techniques in Computer Aided Circuit Analysis." The program understands circuits well enough to determine proper operating voltages and currents by common-sense reasoning, rather than by solving network equations. It can now handle the rather complicated circuits found in IC chips. Specifically, the circuitry of a UA741 op amp has been successfully analyzed. Experts in the electronics field have shown considerable interest in this new, knowledge-based, common sense approach. Sussman believes that the principles will transfer smoothly to other domains and become a general theory of debugging applicable to all engineered systems.

Fahlman and Grossman have done extensive work on a new representational scheme which we believe may have considerable consequences with respect to future machine architectures. Fahlman and Grossman argue for a network of nodes and relationships from which knowledge is retrieved by parallel searches along A-KIND-OF and other links. Their thoughts have been germinating over the past year, but the work was too preliminary to allow commitment to writing until recently. Now both Fahlman and Grossman have completed papers which outline theories, propose experiments, and define, in the case of Grossman, specific application possibilities in the very large data base area. New hardware progress in the integrated circuit industry provides some hope that machines oriented toward such searches may become practical, given that simulation on serial machines proves the ideas sound.

Sandewall has concentrated on the data base problem as well, but with with a view toward dealing with some immediate problems of data base organization and self description. He has worked out a block structured scheme capable of handling both procedures and data, designed to smooth the interface between programs on one side and a collection of data bases on the other. It is expected that his evolving theory will particularly influence our proposed work on very large data bases. Sandewall has finished a prototype system of programs that illustrate his ideas.

Understanding Natural Language

Pratt and his students have concentrated on the problem of dealing with efficiency issues in parsing natural language text. This effort takes us in the comfortable direction of preparing for the time when real application of natural language systems will make a theoretical understanding of the efficiency issue of great practical importance. Pratt's paper of January, 1975, presents an optimality proof demonstrating that LINGOL's parsing algorithm is optimal among those of its class in the sense that all the phrases that it builds up can be used.

Meanwhile Marcus has completed the programming of a new parsing system based on the wait-and-see philosophy for handling ambiguity. Instead of picking one option at all decision points with the risk of being stuck with costly retreat, Marcus' system carries all options forward, making a decision only when the decision can be made with certainty. The design of this parser was finished by the end of 1974; implementation took place during 1975 and is just now complete. The next task is to augment the small grammer Marcus has used in debugging his parser with a much more complete grammar, one large enough to be a good interface to various intelligent terminal or large data base modules.

Progress has also been made on the relatively neglected linguistic problem of generating language. This involves, among other things, a need to know what the listener already knows. McDonald has completed an M.S. Thesis studying the problem. There is no program embodiment of the ideas it puts forth, but creating and experimenting with such a program seems sensible and is planned.

Understanding Vision

Gene Freuder has developed a system for visual recognition, with implications for the general problem of control. The system is intended to provide the flexibility required to deal with realistic visual scenes; he is looking at hammers, not stylized models, but everyday, toolbox hammers, under natural conditions (e.g. lighting, background, orientation, occlusion). He is studying directly the interaction of knowledge and control required in an integrated pass at the entire task, rather than looking at the attendant problems of image processing in detail. Particular results, as they are acquired, combine with general visual knowledge to "suggest" and "advise" further processing. This process, called "active knowledge," extends and implements the principles of "heterarchy" and "domain directed processing" pioneered in this laboratory.

Since description has long been regarded as the key to success with any difficult problem in artificial intelligence, Hollerbach's work on describing curved objects is another important step forward. He has developed an approach towards shape description based on prototype modification and generalized cylinders, a notion invented by Binford at Stanford. His programs describe and identify pottery from vase outlines well enough to be remarkably consistent with the descriptions and identifications given by archeologists. The resulting descriptions seem very natural and suitable for dealing with many sorts of objects besides pottery. The emphasis throughout has been to develop useful, qualitative descriptions which bring out the significant features and subordinate lesser ones.

Marr and his students continue to look at vision from another point of view, that of concentration on thoroughly understanding the problem of translating image arrays into symbolic descriptions without the intervention of high level knowledge. Their thesis is that real image understanding cannot be done without very solid programs working at this level. This basic work seems directly on the critical path leading toward real image understanding, be it of ordinary scenes or ERTS type data. As of January, 1975, the work had concentrated on discovering what features of one-dimensional intensity profiles could reliably provide information about edge types and such modifiers as sharpness of edge focus. Since then, the edge and edge modifier work has been extended to two dimensions with satisfactory results. The most important work, however, has been on the problem of grouping the very lowest level symbolic descriptions of edges together into larger symbolic aggregates and the subsequent realization that a new theory of texture will result from this work. Marr has invented about four new grouping algorithms and prototype programs have been written for them. Some experiments with these prototype programs have been encouraging, but for the experimentation needed, the programs will have to be rewritten for the sake of efficiency. Working with image arrays is time consuming and requires more than the usual attention to program efficiency.

Typical of other work directed by Marr is a theory of visual light source detection developed by Ullman. Several factors were examined, among which were: absolute intensity values, intensity compared with the average illumination in the scene, both local and global contrast, and lightness contrast. A surprising experimental result was that these factors were shown to be insufficient for explaining our ability to detect light sources in the visual field, even under very simple conditions. While high enough contrast, for example, is a sufficient condition for creating the perception of radiance, it was shown that a light source can sometimes be detected in cases where the contrast is very low. A method for detecting light sources was then proposed and implemented, based on the comparison of two ratios:

for a given pair of adjacent areas, both intensities-ratio and gradients-ratio are computed. If neither of the areas is a light source the ratios should be equal. If they are not equal, one of the areas is a light source. The method can thereby detect the source and compute its intensity.

A collection of our previous papers on vision, learning, and representation has just been published by McGraw-Hill, under the title, *The Psychology of Computer Vision*. Sales have been encouraging. This is part of an effort to disseminate our work and speed the technology transfer.

Productivity Technology

Electronic circuit problems were selected as an initial focus for developing automatic manipulation and inspection ideas because the parts are fairly well constrained to a few types and because the vision is somewhat more two dimensional than in most problem areas.

Horn's program for visually aligning IC chips in preparation for lead bonding was a step forward since this job is now done entirely by human eye in spite of attempts made by industry to automate the process. This work was largely completed in September of 1974. Since that time some polishing has taken place, and efforts have been made to transfer the technology to practice. An article on this, "Orienting Silicon Integrated Circuit Chips for Lead Bonding," has been accepted for publication in *Computer Graphics and Image Processing*. GCA, the leading maker of automatic bonding equipment, produces a machine by which one person does the alignment for several bonding stations, but they have not been able to eliminate the human alignment step. The need appears to be particularly acute in the manufacture of hybrid circuits which should not be manufactured by foreign suppliers.

Taenzer has worked on inspection higher up at the level of solder blob checking. Unlike Horn's program, Taenzer's is still too slow for immediate application, but it nevertheless tackles the problems of practical inspection.

Another advance in doing applied vision was reported in Lozano-Perez' work on a system called "PROPAR" which borrows from natural language work in the sense that it generalizes the Augmented Transition Network parser idea in a direction that makes it a workable way of describing real images. Lozano's paper shows that the approach vastly simplifies the recognition of electronic circuit parts which previously was done by cumbersome, opaque, and less powerful programs.

Motion is another area of interest since many applications in both manufacturing and remote navigation require a tracking capability. Speckert has put together a system that tracks a bouncing ping pong ball. He is working on improving both the image handling and the common sense understanding of this system so that it can handle general images with high accelerations.

Manipulation has continued to progress as well. Raibert, for example, has succeeded in writing a program by which an arm learns about the dynamics of trajectories. Because of recently solved hardware problems, his work to date has been confined to three joints, and more work is planned.

Work on force feedback has gone forward using the older, simpler arm by Silver, now improved by the addition of a tilting, force sensitive vice. A radial bearing has been assembled automatically. In the last few months this was complemented by insertion of resistors and IC dips into printed circuit cards. (To be sure, special purpose machines can do this, but at great expense. Moreover, we have done this exercise primarily because issues are addressed which generalize and lead toward assembly problems beyond existing special purpose devices.)

Learning And Debugging

Some debugging issues are addressed by Goldstein's work on a personal assistant scheduler. The idea of debugging is of central importance, despite the apparent difference in the selected domain from others in which debugging is studied. Two versions of a scheduling program have been developed, experimented with, and compared. The first was a simple LISP-based, brute force program with no failure and little debugging; it was implemented to serve as a point of comparison. The second version was implemented in CONNIVER and involved backup along with procedural representations of people and their preferences. The third version, of which many modules are complete, is based on the more advanced frames concept for representing knowledge and involves ongoing implementation of Goldstein's ideas about debugging and bargaining. The version now under study is to work in part by simple first pass approximations followed by sophisticated conflict-resolving debugging specialists. We believe this could lead to a much more accurate and knowledgeable theory of project management since the problem of scheduling in the face of diverse demands and tradeoffs is common to all hard scheduling problems.

Systems Work

We have argued at length that the era of time sharing is eventually going to end because of startling reductions in CPU and central memory costs. To hasten this development, with its corollary reduction in system complexity and software costs, we are working on a so-called LISP machine. Now 99% of the logic design for the "draft" version of the machine is complete and it will be sent out the door for wire wrapping of the boards imminently. Considerable code has been written as well.

To be precise, the system will work by way of two levels of interpretation: LISP expressions will first be compiled into 16 bit, high level "macrocode." The macrocode will then be translated into the LISP machine's particular microcode. Most of the first pass LISP to macrocode compiler is now complete. About 30% of the macrocode to microcode compiler is also done. One version of a complete system is running with simulation of the two compilers running on the PDP10 using PDP10 instructions as the target language instead of the LISP machine's microcode.

During the last six months other system work has included the installation of 256k of new core memory on our PDP10. As expected, dramatic improvement in system performance has resulted and people are once again working with and learning from large programs. Until this memory was installed, really large programs ran so slowly they were beyond practical debugging.

Intelligent Terminals And Very Large Data Bases

Our work on these programs started in January of this year and our progress is still largely in the form of proposals rather than accomplishments. Nevertheless, we have invented and plan to work on several notions we believe to be of fundamental importance to work on the large data base problems, for example, the ideas of semantic paging, semantic cache memory, semantic interleaving of execution and fetching, and the use of the semantic difference net idea for searching.

APPLICATIONS
ORIENTED
BASIC
STUDIES

APPLICATIONS ORIENTED BASIC STUDIES

The Artificial Intelligence Laboratory proposes to work on issues central to the large data base problem, the intelligent terminal problem, and the natural language engineering problem with a broad research program based on the following tenets:

> The key to making computers more helpful and effective lies in how we represent knowledge in them. New representation ideas have exciting prospects. Natural language interface engineering is important. There is a deep need to study common sense reasoning. And people must be understood as well as the files.

THE LARGE DATA BASE PROBLEM

It is now possible, indeed common, to record huge amounts of information. Unfortunately there is no companion ability to use that information intelligently. Large files have become black holes for knowledge in many cases, military equipment maintenance records being a particular example.

Getting useful information out of large files is thus an important practical problem. It raises important theoretical issues as well. Artificial Intelligence has done well in tightly constrained domains -- Winograd, for example, astonished everyone with the expertise of his blocks-world natural language system. Extending this kind of ability to larger worlds has not proved straightforward, however. Even dealing with a story in an elementary reader seems to require a great deal of background knowledge and more sophistication of representation. Handling large technical data bases may be harder still. The time has come to treat the problems involved as central issues.

The Large Data Base Problem And The Intelligent Terminal Problem Overlap

The collection of ideas we propose may be allocated among several headings: large data bases, intelligent terminals, and fundamental artificial intelligence research. We stress, however, that there is considerable overlap both with respect to needs and to potential solutions. To be sure, we feel that knowledge representation and organization form the core of the large data base problem while natural language engineering, common sense reasoning, and user models form the core of the intelligent terminal problem. But we also believe that neither a large data base program nor an intelligent terminal program can be fully successful without substantial dependence on existing and needed artificial intelligence work in both areas. For the moment we dwell on large data bases and on how something like an intelligent terminal system might be used in concert with large data bases. Later on fundamental research is discussed in *ISSUES ORIENTED BASIC STUDIES*

The Small Data Base Problem Suggests Ideas

Before talking about what is difficult, let us begin with what is easy. Some of what we would like to do with large data bases and intelligent terminals has been illustrated already on very simple domains involving only small data bases. The work is important because it serves to establish a context for the discussion of large files. A glance at the better dialogues with Winograd's now classic natural language system shows several important capabilities already demonstrated in miniature.

- *The system has some understanding of natural language* This is essential, not merely desirable for the large file program because establishing an enthusiastic community of users can be easily prevented by just a little too much complexity.

- *The system watches for particular events* It stands ready to alert the user to disturbing combinations. Constantly active "demon" programs watch the flow of information into and out of the data base and complain if appropriate.

- *The system can introspect into some of its own activities; it can answer questions about its own goal structure* This is important because we will often want a large data base system to explain how it arrived at a conclusion or to supply additional arguments to the ones already given.

- *The system translates English descriptions into search programs* This is an important precedent because we believe data bases should eventually include English descriptions of themselves and how they can be accessed and used. If these descriptions can be understood and exploited, the problem of dealing with diverse data bases can be considerably reduced. The English understanding required, being specialized to data bases, may not be much more difficult than that encountered in the blocks world!

A natural question is How much distance lies between these talents of Winograd's system and something that would be useful to a doctor, congressperson, or economist? We believe the distance is considerable in some dimensions and slight in others. In the next section we explain which dimensions seem critical to us and where we believe heavy resource allocation is required.

IDEAS AND SOLUTIONS

New Ideas In Representing Knowledge Have Exciting Prospects

Questions must be answered by finding relevant facts. Finding and relevant are the key words.

We believe that coping with the problems of accessibility and relevancy in large data bases requires understanding of those large data bases in terms of hierarchically organized frames. This is crucial. Creating and using hierarchy in knowledge representations is difficult, and the problem has come to be widely recognized as a central dominant theme firmly entwined with nearly all important artificial intelligence challenges.

Minsky's frame theory has created considerable excitement precisely because it addresses the representation problem squarely and provocatively. Here is the essence of the theory, as explained by Minsky:

> When one encounters a new situation (or makes a substantial change in one's view of a problem), one selects from memory a structure called a frame. This is a remembered framework to be adapted to fit reality by changing details as necessary.
>
> A frame is a data-structure for representing a stereotyped situation like being in a certain kind of living room or going to a child's birthday party. Attached to each frame are several kinds of information. Some of this information is about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed.

We can think of a frame as a network of nodes and relations. The "top levels" of a frame are fixed, and represent things that are always true about the supposed situation. The lower levels have many _terminals_ -- "slots" that must be filled by specific instances or data. Each terminal can specify conditions its assignments must meet. (The assignments themselves are usually smaller "sub-frames.") Simple conditions are specified by _markers_ that might require a terminal assignment to be a person, an object of sufficient value, or a pointer to a sub-frame of a certain type. More complex conditions can specify relations among the things assigned to several terminals.

Collections of related frames are linked together into _frame-systems_. The effects of important actions are mirrored by _transformations_ between the frames of a system. These are used to make certain kinds of calculations economical, to represent changes of emphasis and attention, and to account for the effectiveness of "imagery."

For visual scene analysis, the different frames of a system describe the scene from different viewpoints, and the transformations between one frame and another represent the effects of moving from place to place. For non-visual kinds of frames, the differences between the frames of a system can represent actions, cause-effect relations, or changes in conceptual viewpoint. _Different frames of a system share the same terminals_; this is the critical point that makes it possible to coordinate information gathered from different viewpoints.

Much of the phenomenological power of the theory hinges on the inclusion of expectations and other kinds of presumptions. A _frame's terminals are normally already filled with "default" assignments_. Thus, a frame may contain a great many details whose supposition is not specifically warranted by the situation. These have many uses in representing general information, most likely cases, techniques for by-passing "logic," and ways to make useful generalizations.

The default assignments are attached loosely to their terminals, so that they can be easily displaced by new items that better fit the current situation. They thus can serve also as "variables" or as special cases for "reasoning by example," or as "textbook cases," and often make the use of logical quantifiers unnecessary.

This frame-systems are linked, in turn, by an information retrieval network. When a proposed frame cannot be made to fit reality -- when we cannot find terminal assignments that suitably match its terminal marker conditions -- this network provides a replacement frame. These inter-frame structures make possible other ways to represent knowledge about facts, analogies, and other information useful in understanding.

This theory is a scholarly integration of some old ideas and many new ones. It opens the way to research on previously intractable questions. In particular the problem of establishing a correct context is exposed as a problem to work on rather than a problem to worry about. Minsky integrates and elaborates several approaches in considerable detail:

- The idea of clusters of frames organized like places on a road map around dense local paths, larger and less dense arteries, and finally the limited access superhighways.

- The idea of group characterization by multiple prototypes, d-capitols as he calls them.

- The idea of moving from one hypothesized frame to another by way of difference analysis, the so called similarity net idea.

- The idea of preserving as much information as possible in moving from frame to frame by knowing invariance properties of terminals and subframes.

Elaboration of these is important to the development of entwined representations capable of enabling fast access to large data bases. Fahlman and Grossman are working on other mechanisms which may well prove to be important prerequisites. They are motivated by a feeling that intelligent systems should have a very fast, perhaps even parallel mechanism for working with semantic networks. They argue that questions like these should have nearly instant answers:

> Which ships have bad boilers?
> Are elephants long nosed marsupials?
> Do penguins fly?

Fahlman's progress report profferes a solution in the form of a theory of nodes, links, and propagating labels. The theory was developed after Fahlman tried to apply frames to large problems and stumbled into a fundamental set of problems: symbol mapping (the problem of finding the properties of an object whose description is in the form of a tangled, multi-layered frame); inconsistent fact detection and the enforcement of restrictions; and of finding the right frame during recognition. Falhman's hardware nets, described in an appendix, are one possible solution to these problems which certainly must be solved before frames can be applied to large data bases. As yet the ideas are new to us and we have no solid feeling for whether a Fahlman machine would be a break-through or a disappointment on the order of the resolution-based theorem prover of the past, or the perceptron-based "learning machines." It is clearly important to find out.

We believe that all of the frame accessing ideas will be increasingly necessary (and effective) with increases in data base size. As yet, however, the frame idea is largely untested. Many examples have been suggested to illustrate how frame ideas might be applied, but no working system yet demonstrates their real strength.

We have some evidence that frames can be used to structure large amounts of scenario knowledge of the sort involved in children's stories. Creating such a system for a problem of considerable real world importance is a high priority objective. We have not yet devised frame structures that would capture such scenarios as nations in conflict. We imagine that there could be a frame theory for the participants in international conflict scenarios that would have terminals and defaults for friends, enemies, attitudes, traditions, desires, wealth, and occupations. (Abelson has drafted interesting caricatures of such representations, and leaves open the problem of filling them out with "large data base" detail.) Such a line of work would hopefully lead to systems capable of hypothesizing conflict frames early enough to activate the mechanisms that could help avoid their consequences. Having such structures should allow frame instantiation, for example, from the *New York Times* news summaries, in the same way that birthday party scenarios could be instantiated by stories from children's readers, as suggested in Charniak's work. We hasten to emphasize that dealing with something so complicated as the *New York Times* is very hard and must be considered long range. Yet some of the problems of representing the desires, moods, and likely actions of people and countries are suitable for direct work now.

Frame Structures Suggest Memory Organization Schemes

Ira Goldstein has pointed out that frame theory leads to the notion of semantic cache memory. Selecting a frame by hypothesis, bringing that frame into fast memory, and leaving it there until dislodged by newer frames is good in the same way that ordinary cache memory is good. Once referenced, a frame is likely to be referenced again soon.

It is equally true that frames closely related to a frame being processed are likely to become relevant. Consequently the idea of paging comes to mind. Semantic paging would group together semantically related information units that tend to be accessed together and could be used to bring possibly relevant information up to at least the fast secondary storage level.

In addition, we suspect the frame hierarchy, similarity networks, or other frame linking mechanisms should be used to automatically bring particularly attractive frames out of semantic pages into the semantic cache while processing proceeds. Thus expectations would be used in an overlap of semantic processing and fetching functions analogous to the old idea of overlapping the executing and fetching activities in ordinary machine instruction processing.

Suppose, for example, that a large data base system is asked to supply information about world oil price stability. Dealing with this, the system would follow a thread from the oil frame to the international meeting frame instantiated to specialize it to OPEC gatherings. By virtue of a semantic cache, frames encountered along the way would be retained. By virtue of overlap between semantic processing and fetching, other frames would be brought in dealing with pricing policies, recent geopolitical events, environmental, and other related issues.

Natural Language Interface Engineering Is Important

Representation of knowledge, as said before, is the most vital part of our proposed focus on large data bases. But other topics strongly related to an intelligent terminal focus deserve discussion because in our view they can be studied most effectively by a synergistic program that deals with both large data bases and intelligent terminals in a unified way.

Natural language understanding is an example of a technology that seems vital to both efforts. It is fortunate, therefore, that this is one area where we feel hard work will surely produce good results even though some of the desired scenarios will require considerably more natural language strength than did Winograd's blocks world task. There are now many natural language systems with many good ideas. The field seems ready to gel; it is a time for integration.

It seems clear, incidentally, that natural language should not be limited to accepting questions in English -- to be really useful, the system should accept questions in English, should generate natural English remarks and reports, and should be able to extract information from large English files as well as highly structured files.

In particular we visualize bringing together something like Pratt's LINGOL or Wood's Augmented Transition Net grammar or Marcus' new wait-and-see grammar (described in an appendix) on the sentence fragment level, Filmore's and Martin's case analysis on the sentence level, and Schank's conceptual dependency theory on the semantic frontier.

An analogy with progress in computer vision research may help clarify why we believe rapid progress can be made toward wrapping up work in natural language research at the level of natural language syntax and elementary semantics. On entering a completely new, unexplored area, the first step must be to get a feeling for how to describe things, usually making simplifying assumptions that place interactions in the background. Guzman's early work on analyzing polyhedra is a good example. By looking at many samples, he developed an *ad hoc* theory of how regions should be grouped into bodies. In this early phase, it was appropriate to focus narrowly -- the way shadows interact with other sorts of lines was justly avoided.

During the next stage of research in the drawing analysis work, descriptions were refined. This made the results of analysis more precise, but analysis became more complicated because enrichment means alternatives and alternatives often lead to ambiguity and subsequent search. It is proper and natural in such circumstances to tolerate inefficient search-oriented strategies for using knowledge so as to expose, through experiment, where the key knowledge problems really are. Later on, however, as descriptive apparatus becomes still richer, the interactions become so well understood that alternatives can be explored not through classical search but rather through deferment of all decisions until choices can be made with complete confidence and total commitment. Waltz' work represents this stage in line drawing analysis: first lines are described in a rich language capable of noting lighting characteristics as well as a line's physical cause -- shadow, crack, boundary or whatever; second, constraints among the lines are so well understood that possibilities for individual lines type identification can be carried forward until each untenable choice can be logically eliminated. There is never any backup.

Syntax analysis seems in a healthy state just under the final stage. More and more evidence suggests that the natural ambiguities that occur in language should be handled not by guess and backup but by a wait and decide strategy. This in brief is the thrust of some maturing work by Marcus, an advanced graduate student. His progress report on this subject, included as an appendix, has a comfortable feel. The parser described already handles the following rather complicated sentences in the sense that it separates them into noun groups and establishes the role of each noun group in the case structure.

Yesterday, John gave the book to the girl.

Who did John give all of the books to?

Is there a boy sitting on the bench?

John wanted the boy to be given a book.

The boy John wanted to persuade to sit on the bench hit him.

The system makes the vital structural distinctions between

They are falling rocks.

They are breaking rocks.

which are not at all similar semantically. The system also handles these:

We believed John.

John was believed by us.

It was believed that John was tall.

We believed that John was tall.

We believed John to have been tall.

John was believed to have been tall.

The boy John was believed to have hit gave me a book.

The boy John was believed to have been hit by gave me a book.

Was the man sitting on the bench?

Was the man sitting on the bench walking to the store yesterday?

Is the girl John saw sitting on the bench walking to the store?

Is the girl John saw sitting on the bench?

People Must Be Understood As Well As The Files

Eventually systems must be intelligent not only about the information they contain but also about the people who use them. Like respected assistants, they must know the shape of what we want from brief instructions if they are to supply us with what really concerns us as well as with what we specifically ask for; they should not trouble us with unreasonable demands for describing subject detail or stupid questions about how much we already know. Only if they understand us to that degree can they get the message to Garcia.

Suppose, for example, a known user steps up to a machine and asks for a report on Turkey. That should be enough, in the sense that good default presumptions aimed at knowledge about people, should handle decisions about such questions as these:

- Is the user talking about the bird or the country?

- Where does he like to get information from? The Encyclopaedia? *The New York Times*?

- How much information does he usually want? Two pages? Twenty? An abstract? An annotated bibliography? How is the quantity decision related to recent evidence of interest?

- Presuming the user is talking about the country, in what particularly is he likely to be interested ? Is he normally interested in information about Soviet border issues, U.S. aid, heroin production, the Dardanelles, rugs, Cyprus, or what?

If the user asks to be kept informed in the event of a crisis, then the question of what he is interested in becomes two dimensional: the system must know not only the topics that are to be monitored but also what constitutes a crisis. The fact that some U.S. tourist gets jailed for buying dangerous drugs is not likely to be a crisis to Henry Kissinger but it might be for a local embassy official. The Secretary of State presumably wants to hear about how the Turks are likely to react to changes in the U.S. aid flow, while the fellow who worries about tourist problems probably need not be awakened each time new information on such subjects appears.

Proper monitoring can be handled by development of a theory of user models and a theory of semantic interrupts. For example, such a theory might involve a technique by which interrupts would be generated whenever specific, user dependent frames are activated by tentative hypotheses or confirmed by a complete analysis.

All of the work on "demons" derived from PLANNER and CONNIVER work is relevant, of course, but we need in addition an epistemology of people and what they know. We believe this is strongly related to the other side of the system, the machine, its large data bases, and what they collectively know. In particular, we believe it will be possible to use the same frame oriented knowledge representation theory for organizing knowledge about both people and the domains that interest them. It should be very productive to keep this in mind, especially since the ultimate system cannot be had without significant progress on modeling both, however similar or disparate they may be.

Naturally users should have easy, smooth mechanisms for resetting the knowledge the system uses to decide when to bother them. Probably most users will require something between absolute importance threshholding, and constant information rate, independent of how much action there is --

something like what we get on the financial pages.

Some relevant work is already underway to handle the problem of people and their preferences in the personal assistant context. After all, even to arrange a person's schedule some facts about when he likes to get up, whom he likes to see, when he can tolerate interruptions, how important he is relative to various superiors and employees, and the like must be known and used. As additional modules are added, particularly in the dimension of automatic note digestion and information retrieval, more progress in this direction will be needed.

For a major large file program, however, additional professional help seems of preeminent importance. In this we are fortunate to work in close collaboration with people having deep interest and considerable experience in thinking about how people think and what they know. We can consequently propose a broad approach, one dealing both with the artificial intelligence issues and some of the human modeling questions. Indeed we believe that a synergistic approach based on such collaboration is far more likely to be productive than separate, divided assaults.

Several Domains Seem Suitable For The Basic Research Work

We want to look at the activities of various professionals including the following:

- The Equipment Maintenance Officers. This of course is the prototypical large file user. Direct attacks on existing files may cloud some issues at the start, but as a domain for study it is interesting by virtue of the potential size of the data base, the need for trouble oriented watchdog demons, the possibility of using common sense qualitative reasoning, and the need to maintain sensitivity to user suggestions about just how to watch and monitor things.

- The Strategic Materials Analysists. The idea is to worry effectively about reserves of various metals and other strategic materials. The problem features both hard tabular data and English phrased opinions about political trends.

- The Logistics Planners. The vast problem of getting things where they should be and tracing blunders derserves careful attention. Part of the problem is a resource allocation problem with conflicting goals.

Again we stress our opinion that all of these areas have much in common in terms of the scientific challenges they offer.

SOME QUESTIONS

Why Do We Concentrate On Text Files?

We do not, really. We believe that the text examples we have given cover the main issues of semantic organization that all large data base problems present. Certainly considerable thought will be given to more structured files as well. We favor selection of target domains in which combinations of text and structured data naturally occur. An example would be maintenance records in which free text and highly constrained forms are blended together.

How Large Is Large?

We believe the hierarchically structured, frame oriented approach we propose will lead to systems that perform well on large data bases whatever the size. Larger data bases will require more muscle in the knowledge structure, but not in linear proportion. By adding a layer or two of depth to a hierarchically organized data base, huge increases in the size of the information covered can be achieved using basically the same theory used at higher levels. The main obstacle then becomes that of understanding the material fast enough to prepare the lowest level description frames as it goes flying by.

At this time so little has been done we cannot realistically speculate on the size of the data bases that will be handled in the near and middle term. But since the knowledge organization research we propose seems fundamental, we feel it will be central to handling any data bases that can be handled at all.

THE PROPOSAL AND THE MILESTONES

It is clear that work on large files interacts very strongly with other foci of laboratory interest. While it is not possible or appropriate to draw sharp boundary lines into the picture, it is possible, however, to guide both new and existing resources toward a large file focus. This point of view is reflected in the following plan elements. Target dates range from 6 months for creation of a structured data base with natural language input to 30 months for implementation of semantic paging and similar ideas.

- Implement a structured data base with natural language access for a small subset of the relevant statistics for a particular strategic material or equipment maintenance data base. This represents a capability already achieved for other, smaller domains. Accomplishment of our goal here will be marked by being able to direct simple statistical queries in English to the data base.

- Concentrate our natural language effort in the intersection of the problems raised by our personal assistant focus and large file issues. After two years time we would expect good natural language front ends to a variety of personal assistant and large data base modules.

- Push forward simultaneously with the network taxonomy work of Fahlman, expecting a Ph.D. in June of 1976. This is to contain ideas and experimental results defining the possible role of a fast, well conceived A-KIND-OF processor in the large file program.

- Implement a frame system representing interpretive knowledge about the data base. This interpretive knowledge consists of both common sense facts such as the fact that oil must be refined as well as expert knowledge such as the time and costs involved. This milestone represents the application of existing theoretical ideas to an application area. This allows queries to go beyond simple statistical inquires -- data bases can be asked about trends, possible dangers, etc.

- Implement a demon capability for real-time monitoring of critical situations. Demons will monitor both symbolic and statistical additions to data bases. We must solve technical problems which come up when demons are activated by complex patterns, and we must solve problems of computational efficiency needed when multiple demons are activated.

- Extend the natural language capability so that a system can add statistics to a data base by processing text. The system is freed from the time consuming and error prone process of human reading of textual data and human entry of the relevant facts. Furthermore facts can be added more quickly and the information flow can have less time lag. This of course is a critical issue for data bases representing estimates of the supply of critical defense items. This milestone will be achieved when a system reads news articles on a sample material.

- Use of frame organization and especially hierarchies and differential frame pointers to allow semantic paging, semantic hashing, semantic pattern matching. It is this organization which goes beyond traditional ideas that we expect to allow reasonable computation times even on very large data bases.

- Keep these diverse research activities coordinated and integrate the results into a pilot system as rapidly as the ideas reach a sufficient stage of maturity. This would be the duty of staff researchers commencing work in early 1976. The result is to be a well documented system continuously available for experimental use as several of the natural language systems are now.

HARDWARE

It is conceivable that a machine inspired by Fahlman's research on knowledge structure might become part of a successful approach to certain aspects of the large knowledge base problem. In our opinion, the Fahlman concept is the first hardware design that promises to handle information problems for which the classic serial Von Neuman architecture may ultimately prove inadequate as the total foundation on which to build a truly intelligent system. We do not propose to begin the design of such a machine now because the idea is not sufficiently mature and it is not sufficiently substantiated by simulation. After this work is done, however, we hope it will prove that the design of a machine should be in a future proposal.

Meanwhile it is clear that developments in artificial intelligence problem areas will require very heavy demands on computing resources. Experience shows that very large address spaces and dedicated central processors are required to cope with the complexity barrier. At the moment the LISP machine concept, proposed for development in a separate section, seems well suited to this purpose.

TECHNOLOGY TRANSFER

We plan to work with a view toward providing implemented modules with both the enchantment appeal and the inherent usefulness required for export. We further plan early involvement of people from a selected user community, and we realize that the availability of such involved users is an important consideration in domain selection.

PERSONNEL

At least the following people would contribute to a large file effort. Negotiations are underway to recruit other artificial intelligence specialists from outside the M.I.T. community.

- Professor IRA GOLDSTEIN will be the project coordinator. He has been leading the laboratory's personal assistant work aimed toward understanding the scientific questions involved in the development of a system capable of scheduling appointments and events, organizing notes and observations, and retrieving a wide variety of knowledge. FRED KERN is currently helping Goldstein with the scheduler module.

- Professor MARVIN MINSKY will work on the problems of knowledge representation and common sense reasoning. It is expected that Minsky's frame theory will be of considerable importance. In studying the problems he will be assisted by several graduate students.

- SCOTT FAHLMAN and RICK GROSSMAN both came independently to the view that some better approach to retrieving knowledge is demanded by the current problems facing artificial intelligence. They are working hard on the new network label propagation idea to handle the sorts of scenarios we see coming up in the large file context.

- Professor SEYMOUR PAPERT is interested in the question of organizations, the managers that exist within them, and the problem of supplying those managers with information that is at once useful and desired. Such questions will certainly be important in the intelligent terminal context since one function of an intelligent terminal would be to facilitate the interaction of an entire office full of people. STEVE ROSENBERG and JIM STANSFIELD will assist Papert in helping to identify the genuine management information issues and use these in the process of selecting a target domain with which to work.

- Professor VAUGHAN PRATT and his students are moving forward rapidly with the development of LINGOL, a very efficient natural language system facilitating grammar writing. This is expected to compete favorably with Wood's Augmented Transition Network formalism for the prize of being the best simple system for implementing natural language front ends.

- CANDACE BULLWINKLE, DAVID McDONALD, and MITCH MARCUS are graduate students specializing in natural language studies. Bullwinkle is working on a module for interfacing various personal assistant facilities to users, McDonald is working on knowledge based dialog, and Marcus is finishing a powerful new parser believed to integrate the best features of a number of separately developed systems. All together we see this work moving toward an understanding of the relationship between task size and complexity on one hand and system sophistication on the other. Natural language interface engineering seems plausible now.

APPENDICES

FAHLMAN: A SYSTEM FOR REPRESENTING AND USING REAL WORLD KNOWLEDGE

MARCUS: WAIT-AND-SEE STRATEGIES FOR PARSING NATURAL LANGUAGE

APPLICATIONS
ORIENTED
VISION
AND
MANIPULATION

APPLICATIONS ORIENTED VISION AND MANIPULATION

The Artificial Intelligence Laboratory proposes to work on applications of computer vision and manipulation with the emphasis shifting from tool building toward solving real problems, using well-known, commercially important problems as a source of challenges. In particular, we plan to devote part of our effort to a hard assembly problem involving putting together or taking apart devices with flexible components, awkward geometries, and poorly constrained tolerances. We further plan to broaden our activities by inclusion of work on tracking moving objects and on real-world inspection, these problems now being within reach of solution. Success on these new problems will lead naturally and smoothly to progress on some aspects of the general problem of applied image understanding, although not all of them as we argue later when we discuss the future of basic studies in vision.

Productivity Technology Has An Important Mission

Broadly speaking, the overall goal of our work in applied vision and manipulation is to develop uses for artificial intelligence techniques in hand or within reach now. We have this goal because we want to do the following things:

- Maintain the economic strength of the country in the competitive international arena through automated production, particularly in assembly.

- Create a more flesible brand of manufacturing capable of very rapid response in times of rapidly changing need.

- Prepare for automated mining, both underground and undersea, as well as automated farming, space exploration and recovery, ordinary maintenance, and maintenance in hazardous environments. In part this means creating an autonomous vehical guidance technology capable of coping with unprepared environments.

We believe these objectives are vitally important because a preeminent computer science and industry is one of the major strengths of the United States. We must translate this advantage into corresponding capabilities in areas where our world position is eroding, especially in productivity. We think changes will come about quickly when a certain threshold of technology is passed over. Just as pocket calculators swept the country like a tidal wave, we believe computer vision and manipulation will soon sweep through the manufacturing and maintenance industries. But before discussing what *soon* means, let us look at the path on which we have been on and how it reaches the stated objectives.

Progress In Transferring Artificial Intelligence Ideas To Industrial Practice Involves Four Major Steps

We measure progress in terms of where we are in executing the following plan:

1 Pick the simplest possible domain, do vision and manipulation in that domain, and use the results to measure the difficulty of the problems in order to allocate resources and predict when applied results should emerge.

This has been done, of course. The blocks world was the domain; the difficulty of vision dictated the concentration of effort on it and the relative neglect of manipulation. It became clear that the existence of global knowledge and the non-linearity of noise effects render most previous work on image processing of little value. Instead, problem solving, constraint exploitation, and symbolic low level processing became central issues as discussed elsewhere in our discussion of directions for basic studies in vision. It also became clear that vision is very hard, and for the sake of near term application results, the goals in vision research should be divided into no-tricks basic study and no-holds-barred, special-purpose, domain-dependent work.

2 Once the problems are uncovered through suffering in the trenches with the "simple" domain, then it is time to get together hardware and software tools capable of meaningful, cost conscious prototype arrangement to solve particular problems.

This too is behind us. Our so-called Microautomation Laboratory exists and supports applications work. It is described in brief later on.

3 Use the tools to tackle real problems. At first these may be selected because they are amenable to solution through the tools and techniques in hand -- obvious undisputable commercial viability is a secondary objective.

Our work on electronic assembly was chosen for suitability in this phase. Many results, described later, are in hand. We believe we will complete the remaining work in electronics this year.

4 Finally, unleash the technology on the most serious problems where the saving would be greatest in human and fiscal terms.

This is to be our primary purpose during the contract period. At the moment, this purpose translates to the assembly of mechanical devices exhibiting all the tough assembly problems we can find which seem essential to industry and apropriate for solution by some combination of vision, force sensing manipulation, and now understood Artificial Intelligence programming ideas.

- Industrial exposure is of course vital to a program like this where sensible resource allocation and technology transfer are so important. We are therefore fortunate at MIT to have a very active Industrial Liason Office which has helped arrange for us a wide spectrum of contacts in industry.

Production facilities visited include those of Texas Instruments, Data General, Digital Equipment, IBM, the RCA hybrid circuit facility, the Kennecott Copper Company Peabody coal mine, General Motors car assembly plants, Delco integrated circuit facilities, AMP incorporated electrical connection and electrical assembly facilities, the Foxboro industrial controls facility, United Shoe Company's automated electronic equipment machinery production, and the GCA facility charged with the production of advanced lead bonding equipment. More of this industrial contact is planned. In particular a number of advanced production facilities and research laboratories will be visited in Europe this fall, again to be financed by the MIT Industrial Liason Office.

We will return to the subject of hard problems in a moment.

THE FOUNDATION FOR THE FUTURE

We now give a brief description of the laboratory, move then to a brief discussion of some progress made with it, and then finish with the proposals for the future.

In our last major proposal, we emphasized two opinions:

- We argued a need for a workable, inexpensive set of hardware and software tools for doing research in applications oriented computer vision and manipulation. We laid out a plan for the creation of these tools.

- We suggested that tool creation without one or two target domains can be inefficient and leads to poor resource allocation. We specified the problem of electronic circuit card manufacture and repair as suitable for bringing general issues into crisp, manageable form.

All of this work is well along as evidenced by our progress report for 1975. The inexpensive, exportable vision and manipulation laboratory now exists. It has demonstrated its usefulness by supporting important target domain projects involving electronic circuit cards and by facilitating first steps outside electronics in other worlds of interest like the world of small mechanical devices.

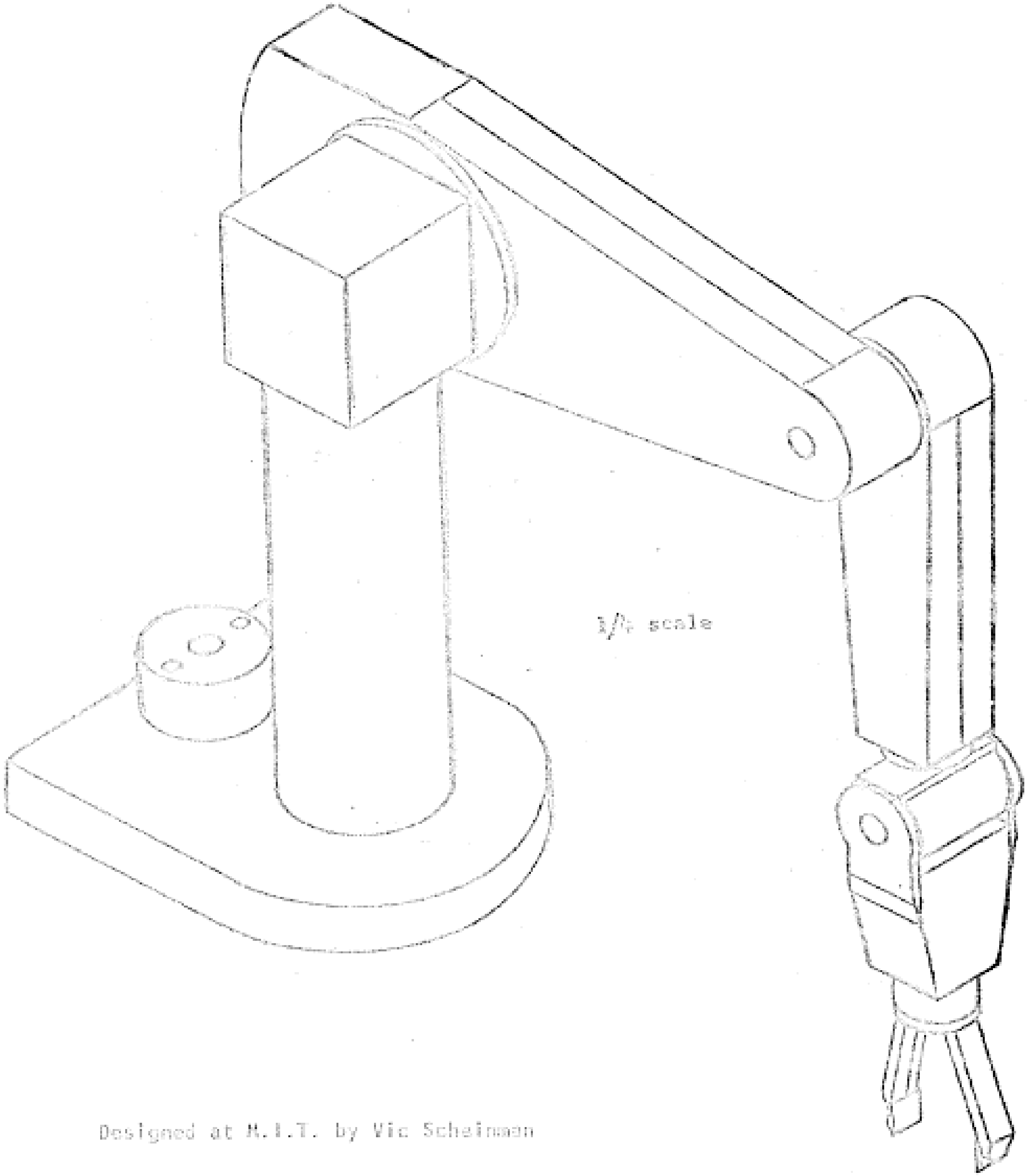
The Hardware Is Ready

We have a PDP11/40 with extended instruction set (EIS), 31K of core memory, two 1.2 megaword disk drives, and a GT40 graphics display terminal. In addition, we have the following devices:

- Two VICARM mechanical arms, 3/4 human size, with six joints and six degrees of freedom, position and velocity sensors in most joints, and gripper hands. These comprise our primary manipulation system. Texas Instruments, the University of Illinois, and the Naval Research Laboratory have ordered the MIT designed arm and more are expected to follow suit. Figure 1 shows the general appearance of these arms.

- A Vidicon picture digitizer, consisting of a television camera and a television monitor, and a digitizer which converts the scanned picture to 8 bit digital data at a maximum rate of 65 microseconds for each point in the TV scan. This is currently our primary vision system.

- A Hughes modulatable laser with an output of .001 watt, used principally as a pointer or tracking device.



Designed at M.I.T. by Vic Scheinman

- Two motor-driven mirror-deflection systems, by Spatial Data Systems, Inc. Each mirror system has two independently controllable mirrors with perpendicular axes of rotation, and settling times of approximately 3 milliseconds for motion of the mirrors. One system is used to deflect the Hughes laser; the other is attached to a PIN diode via an optical system and is used for scanning scenes as a slow but very linear, high resolution vision input device.

- An Optronics International, Inc. photowriter, capable of making extremely high resolution photographic negatives on film from stored digital pictures.

- A precision X-Y table, made by the Icon Company, capable of motions in two orthogonal directions with an accuracy of .001 inch, over a range of motion of 6 inches with a maximum velocity of 4 inches/second. This device is used to position work under the mechanical arm or the vision system.

- An Analogic digital-to-analog and analog-to-digital converter with 8 D to A channels and 64 A to D channels, capable of performing data conversions in 3 microseconds, and to which we have added an additional 16 D to A channels.

The Software Is Ready

The Productivity Technology Group PDP11/40 uses standard DEC software, the Disk Operating System (DOS), in an effort to make software developed within the group useable elsewhere. The group has developed the following software:

- VIDIN is the visual input program. It is used in conjunction with either the Vidicon or the PIN-diode/mirror system to obtain digitized pictures which are stored on disk. The user can select arbitrary sub-sections of the entire picture being scanned, and the program continually displays to him the sub-section of the picture which he has selected.

- MAPPER is a program for computing statistics on stored digitized pictures. MAPPER can provide histograms and intensity gradients, printouts of the raw data and even graphic representations of the picture itself. Operating in conjuction with a program that runs on the GT40 graphic display terminal, it can threshhold a picture into 8 intensity gradations, and then cause the picture to be displayed on the GT40. MAPPER enables the user to examine small subsections of the entire digitized picture for detail, or to combine subsections together to get a larger view of the statistical analysis of the picture.

- LISP, a fully-general LISP interpreter, with shallow bindings and a non-recursive garbage collector which conserves memory. This LISP does have the restriction that atom print-names must be 3 characters long, or less. It does standard DOS file I/O, and can additionally do input or output from a variety of terminal devices, or from the Artificial Intelligence Lab's PDP10 timesharing system.

- RUG, a symbolic, DOS-oriented debugger, which replaces the DOS debugger ODT. Using RUG, a user may debug standard DOS format programs; the debugger makes available to him all his symbols, disassembles the contents of core into assembly language statements with symbolic references included, and allows him to enter assembly language code into core by merely typing it in. RUG has various numeric and special-purpose type-in and type-out routines (such as ASCII and RADIX-50). It allows the user to set breakpoints in his program, and additionally allows him to monitor the contents of a core location, register, or device for some desired condition.

- A set of arm control routines for the VICARM mechanical arm. These are implemented in LISP, and include routines for moving the arm to a particular state-vector (essentially a position and velocity for each joint), and for halting the arm's motion and holding it still. The routines incorporate a trajectory planner which decides on the route the arm will take to get to the desired state-vector, and a dynamic routine. The dynamic routine looks at the desired state-vector as computed by the trajectory planner at the end of the next time increment, and computes the torques which must be applied to each joint in order to achieve that state-vector.

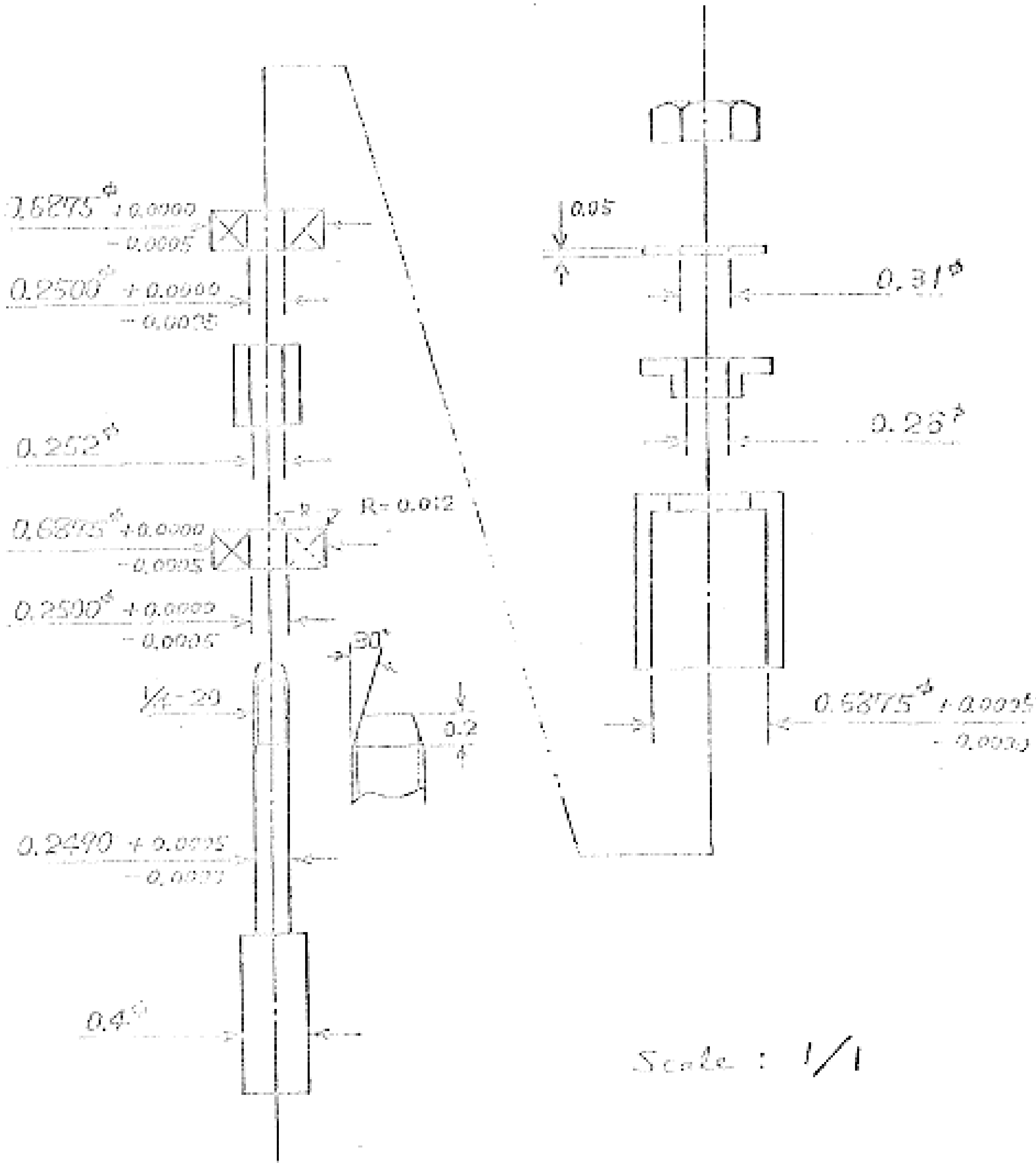
Finally, there is an improved editor which utilizes the graphic capabilities of the GT40 terminal to implement a real-time edit mode, and there are programs for transmiting files to and from the timesharing system, for displaying images on our GT40 termininal, and for other typical system activities.

Artificial Intelligence Research Feeds Productivity Technology Development

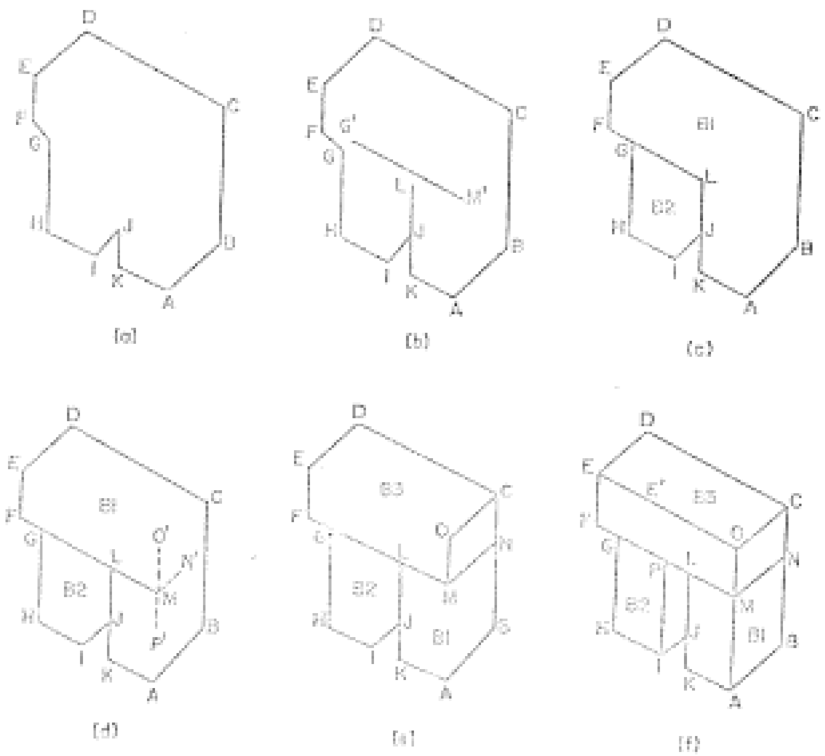
The flow of basic Artificial Intelligence techniques to applied problems has begun in the Artificial Intelligence laboratories and a complete list of achievements in this direction by the community would surely include the impressive work done in Japan on A.I. inspired ZIP code readers and circuit card inspectors, the work at Stanford on manipulator control, and our own work, of which the following are of particular relevance:

- Inoue and Silver demonstrated the automatic assembly of a radial bearing with 12 micrometer tolerances, thereby demonstrating that *complicated assemblies can be made using flexible, programmed manipulators*. The achievement further demonstrated that *knowledge based force sensitive manipulators are the key to advanced assembly operations*, and that *high level programming languages make programming for mechanical assembly a viable concept*. The bearing is illustrated in figure 2.

- Horn showed how to apply the technique of *knowledge based profile analysis*, borrowed from earlier blocks world work by Binford, Shirai and others, to the applied problem of orienting IC chips in preparation for lead bonding. Figure 3 shows some typical profiles encountered. Mat Mason is doing the same thing with automated diagnosis of electronic signals of the sort that technicians look at on oscilloscopes.



Size and clearance of the parts.



Steps in analysis.

- Taenzer uses *region growing ideas* to inspect solder joints for the standard defects -- pin holes, cold joints, and open holes. See figure 4.

- Speckert has begun to incorperate *higher level knowledge* about obstacles in low level tracking routines that follow objects through otherwise difficult accelerations. Figure 5 shows a trace of Speckert's points of observation as his system tracks a ping-pong ball through a bounce.

THE CHALLENGES TODAY

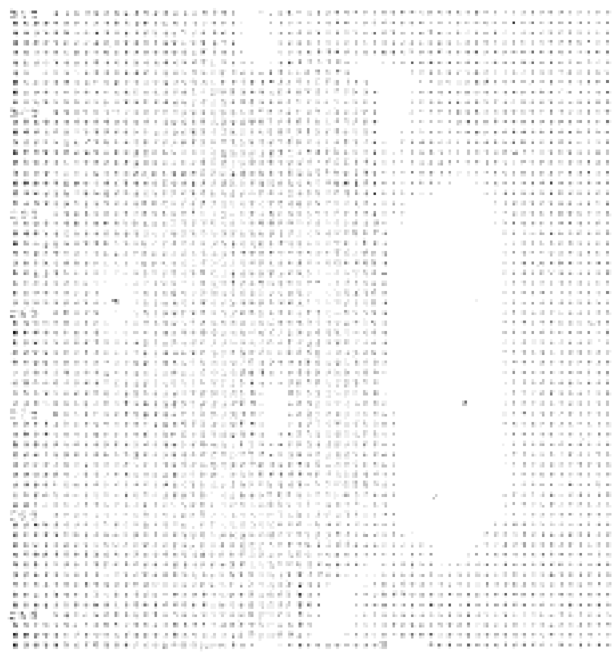
Using our now well developed tools, we plan to accomplish the following:

- We plan to conclude our ARPA sponsored automatic assembly program. This will be amplified in a moment.

- We plan to deal with a difficult, important, and costly inspection problem, solve it, and use the resulting solution to determine if cost effective solution is in sight. We propose the inspection of jet engine castings as the candidate domain.

- We plan a new program in the dynamic vision of tracking and vehicle navigation. We believe this will prove to be a fine source of new important problems suitable for replacing our work on assembly as that phase of our program goes to successful completion.

Solder joint with pin hole



LOCUS OF POINTS SCANNED WHILE VISUALLY TRACKING
A BOUNCING PING PONG BALL

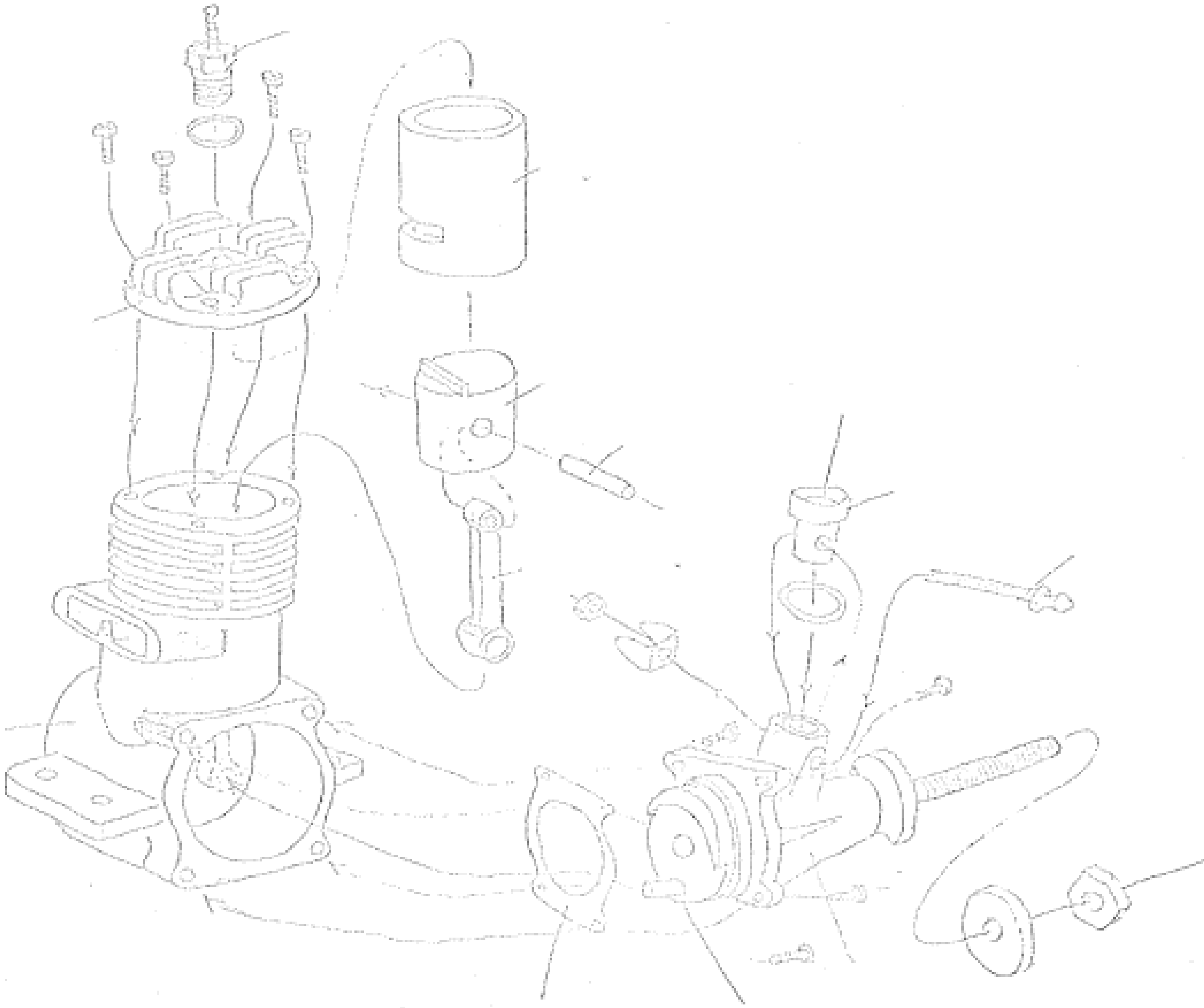
Assembly Of A Gasoline Engine Addresses The Tough Issues

It seems to us that we should be working on elements from the overall list of problems that are best fitted for us in terms of our own expertise and in terms of what workers at other institutions are doing well enough. We therefore propose to build the small gasoline engine -- very likely the model airplane engine illustrated in figure 6 -- with a view toward solving the following problems:

- The programming problem. Our intention is that the assembly program should look like an annotated parts list with advice about the assembly, not low level instructions. We are inspired in this effort by complementary work at Stanford and by our own history of success with very high level languages like PLANNER and CONNIVER.

- The parts orienting problem. The parts are to be scattered at random on the table or drawn from bins. The SRI results will be adopted, if sufficient.

- The flexible part problem. The engine has a gasket. We hope to connect up the fuel line as well. We believe our handling of flexible wires in our electronics assembly work will be valuable experience.



- The two-hand coordination problem. The engine has a difficult problem inasmuch as piston, piston pin, and connecting rod must be brought together. Inoue's and Silver's work with the radial bearing, presented in the progress report, gives us some solid close-tolerance manipulation experience.

- The straight coordinated movement problem. In an effort to eliminate the complex equations normally required for dealing with arm dynamics, Raibert is constructing a model whereby the arm controller learns from previous attempts at movement control and stores this learned behavior into a large data base on disk. The first step is to cause slow motion of the arm between pairs of points in order to teach the system the gravitational parameters along the path in between. Then the arm controller can begin to learn about the dynamics of motion along the same paths by iterating the motions at speed, trying to learn about the arms dynamics in order to reduce the gap between the previous iteration and the desired motion.

- Coordinated vision and manipulation. Insertion of the crankshaft into the connecting rod offers this challenge. The try will not win every time, we suspect, bringing in the opportunity to deal with fault detection and recovery in a natural way. Speckert's tracking work may be useful here, as well as Taenzer's inspection and insertion work.

- Tools. The entire job will require some serious attention to thought about the manipulator's tools.

- Inspection. We may, if fortunate, be able to test the assembled engine and even adjust the needle valve. Inspection for surface defects, however, is to be done in the casting domain, described later.

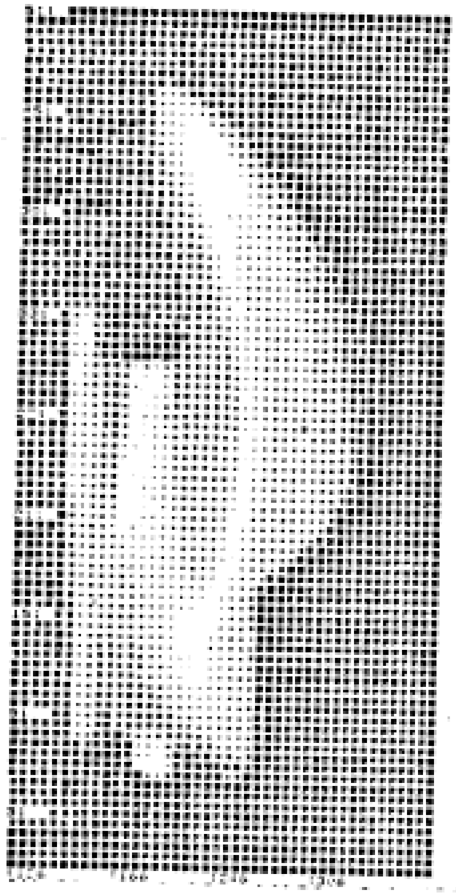
Castings Seem To Be The Right Domain For Visual Inspection Studies

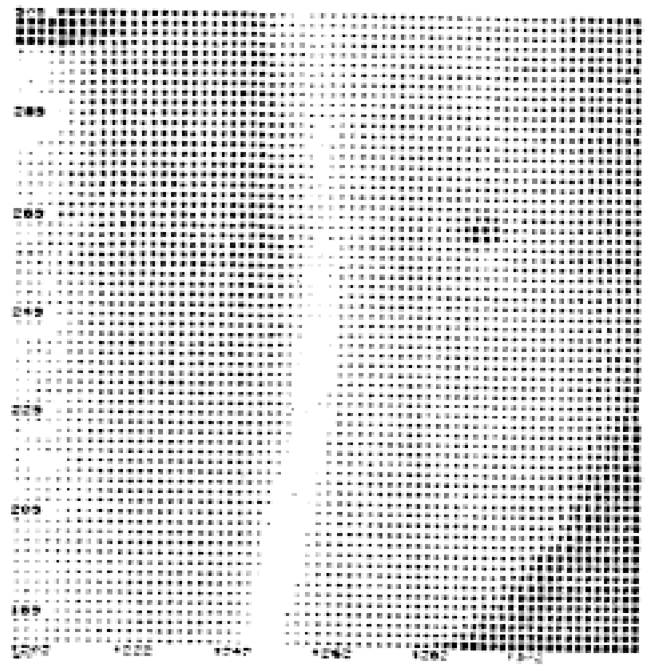
The visual inspection of castings is to be a step toward the development of practical visual inspection systems. Figure 7 illustrates the sort of image produced by a pinhole defect, one of many types to be studied. In our laboratory, we have been developing the technical tools needed to understand images of complex surfaces like these. The current work of Marr, in particular, offers insight into what to do because it deals expertly with the problem of texture.

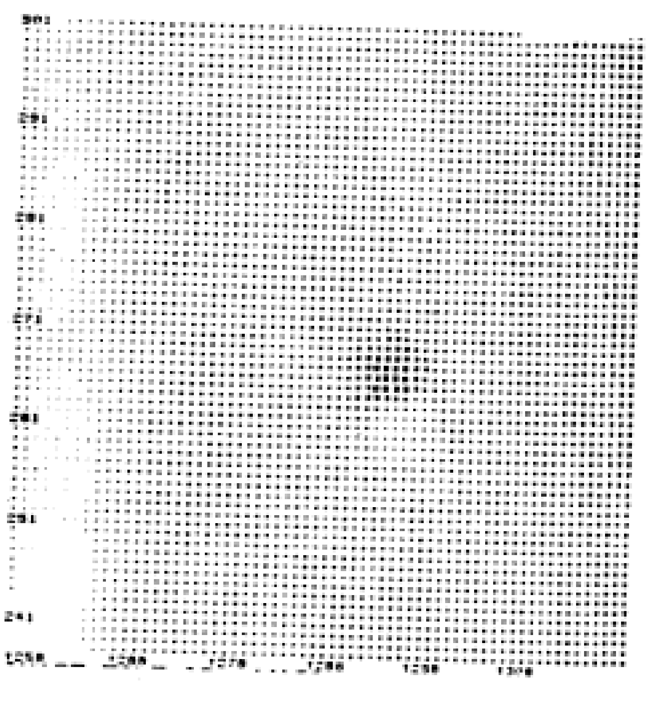
In order to become aware of the practical issues associated with the inspection of metal castings, Woodham has visited the Advanced Casting Laboratory of Pratt & Whitney Aircraft and the Draper Division foundry of Rockwell International. They have been extremely interested and helpful. They have been not only a ready source of defective castings but they have also provided Woodham with the detailed specifications by which various parts are currently inspected. These inspection standards provide the appropriate benchmark for assessing the performance capability of visual inspection systems.

The Dynamic Vision Of Self-Orienting And Tracking Is Ready For Study

Over the years, the problem of seeing and understanding static scenes proved so hard that anticipated work on the vision of motion remained out at pasture with few important results emerging from the isolated attempts that







were made. While there are still many difficulties, the time has come to look at motion problems seriously because we have better ideas about representing visual knowledge, because we have better hardware and software tools, and because the need plainly exists.

To this end, we plan two efforts. One now underway as a doctoral thesis project attempts the following scenario:

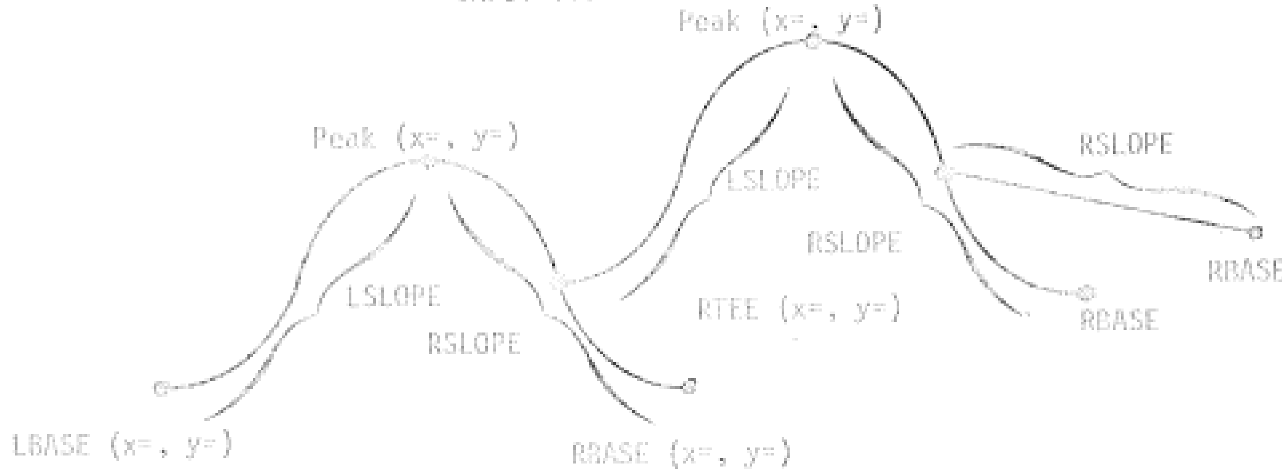
- A vehicle is moving on or over hilly Mars-like terrain. The vehicle wants to keep a good model of the ground it is moving over with a view toward tracking its position on a map if it has one or toward making a map if it does not.

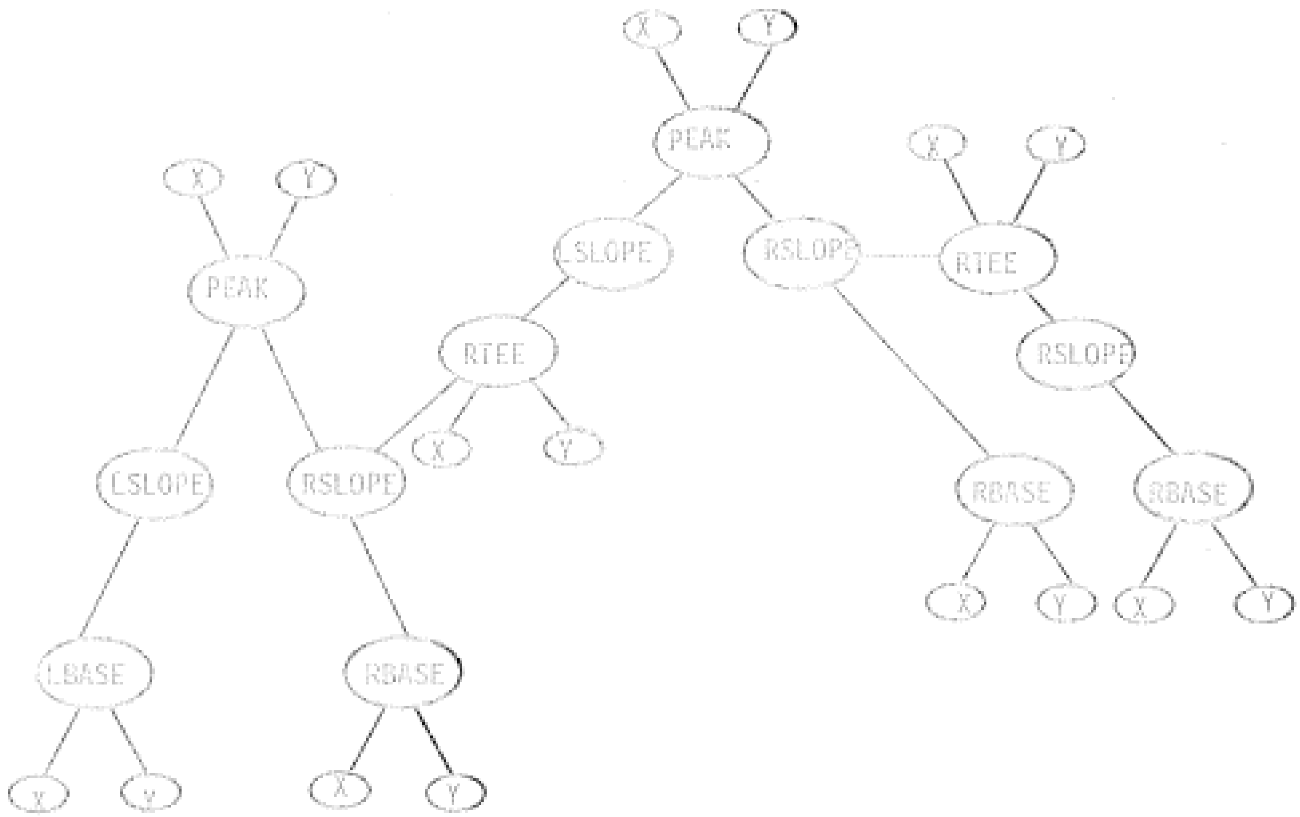
The self orienting study will address two key scientific issues: first the issue of dealing with representing situations evolving through time; second the issue of exploiting constraints evidenced only by comparison of a sequence of situations, not by single situations alone.

As the system looks at the world, it is to make a symbolic network description of what it sees consisting of nodes representing hills, hill features, and the T-joints that appear where one hill obscures the background boundary of another hill further behind. The nodes are linked by relational pointers with labels like IN-FRONT-OF, TALLER-THAN, PART-OF, and so on. Figure 8 illustrates. The changes in these descriptions through time, it is believed, will be subject to handling in terms of Minsky's theory of frames, the now maturing theory of network description, defaults, contexts, expectations, symbolic transforms, subgrouping, and more, as described in *APPLICATIONS ORIENTED BASIC STUDIES*. Each symbolic description of one instant in time will be one frame in a collection.

As the system composes its frames, it will use constraints devolving from motion parallax and from changes in the observed vertices. By so doing, the motion becomes a positive contribution to analysis, rather than a

INPUT VIEW





SYMBOLIC NETWORK REPRESENTING ABOVE INPUT VIEW complication to be removed. This aspect of the work has strong antecedents in Waltz' work on static scene analysis which showed that by properly addressing the constraint question, shadows and the influence of one vertex on its neighbors become terrific aids to analysis, not handicaps.

Another problem, to some extent the inverse of self orienting, is that of tracking moving objects:

- Constrained objects are moving through an image. Knowledge of the constraint is to be used to help track them through situations involving temporary interruptions of vision, as when one object obscures another; and involving fast changes, as when a ping pong ball bounces off of a paddle; and involving objects constrained to move together, as when vehicles follow one another.

The study in tracking will build on the work of Speckert appearing in the progress report. Speckert has tracked ping-pong balls through bounces which would baffle most tracking algorithms by combining simple world knowledge with ordinary adaptive mechanisms. More of this will lend itself to trackers good enough to, say, track people or to allow vehicles to follow one another in convoy, as in a mine or on Mars.

TECHNOLOGY TRANSFER

We plan several steps which we believe will considerably shorten the time it would otherwise take to bring our vision and manipulation work into practice on the factory floor. The work seems too important to depend on the slow osmosis that would occur without deliberate planning. For the most part, these transfer plans involve spinoffs from our normal academic duties which are done with no cost to our research contract.

- We have introduced a new MIT course on the subject of computer vision and manipulation in order to prepare new graduates for active leadership in the field. A text has been undertaken as a mechanism for encouraging such courses elsewhere.

- We have introduced a summer session course for the MIT industry oriented summer program. There was a favorable first-year turnout with engineers and managers from Boeing, Honeywell, New York Telephone, The U. S. Department of Transportation, Westinghouse, Kodak, and Xerox attending. General Motors has purchased our Videotape introductory Artificial Intelligence course.

- We plan articles for the trade literature describing the laboratory equipment and summarizing what has been done with it so far. This is to complement our normal publications in the scholarly journals.

- We have whenever possible built the vision and manipulation laboratory out of easily obtainable equipment and we have avoided modifying this equipment. When new equipment has had to be designed, we have avoided one-of-a-kind in-house construction.

PERSONNEL

- Vision and manipulation work will be done under the general direction of Professor BERTHOLD K. P. HORN, well known for his work on machine vision, including work on shape from shading and reflectivity analysis.

- MEYER BILLMERS will be in charge of system software development. JOHN PURBRICK will do mechanical design.

- DAVID SILVER and TOMAS LOZANO-PEREZ are working with DAVID TAENZER, and MARK RAIBERT toward achievement of the engine assembly.

- BOB WOODHAM will work on the visual inspection of certain castings. MARK LAVIN will do the dynamic vision of self-orientation. GLEN SPECKERT will complement that work with his knowledge-based tracking effort.

ISSUES
ORIENTED
BASIC
STUDIES

ISSUES ORIENTED BASIC STUDIES

The Artificial Intelligence Laboratory proposes to continue a basic studies program aimed at dealing with deep, central issues. The issues are to be exposed by work in computer vision, common sense reasoning, and expert problem solving. The ubiquitous representation problem is involved in all three areas.

THE NEED FOR BASIC STUDIES

All research lies on a spectrum between pure development and pure research. Most work, including all of ours, lies far from either of these extremes. Still, one can either select a need and bring tools to bear on it or one can select issues and choose the best domains in which to expose them. The work described in this section lies more toward the issues oriented focus, where it is harder to specify timetables or even to predict accurately what will happen. Historically some of our laboratory's greatest contributions have come out of work like this that can be aimed but not tightly steered. But this is not to say that our basic research is done by guess and lurch. Our long range goals are clear. They are described in sections devoted to machine vision, common sense reasoning, and expert problem solving.

All three of these sections emphasize the importance of sound representations. *We believe that proper representation is the key to advanced vision, common sense reasoning, and expert problem solving, just as it is to many other aspects of Artificial Intelligence.* Indeed a subtitle to this section, *ISSUES OREIENTED BASIC STUDIES*, could be, *ASPECTS OF THE REPRESENTATION PROBLEM*.

SEEING AND UNDERSTANDING WHAT IS SEEN

The long range overall goal of basic studies in vision is the creation of human level ability in computers that can understand, describe, and search for things in images. To the extent that processing visual information requires intelligence, our studies of vision help us to understand more general issues. We have and will continue to learn a lot about problem solving in general by studying vision problems in particular.

It must be understood that the overall goal of creating human level visual competence has proved difficult. At first the critical path subgoal seemed to be the very problem solving issues we now know much about. Now we feel, on the basis of considerable experience, that the key is the extraction of as much information as possible using relatively knowledge-free techniques. The purpose of this section is to defend our point of view and to break the critical path subgoal of image-to-symbolic-description translation into milestones and forecasted dates.

Creating Computer Vision Is Hard

The study of computer vision began very early in the history of artificial intelligence. Originally the problem of seeing and understanding what is seen was to be a testbed for working out deep problem solving and program organization issues. Eventually this testbed purpose was served, but only after overcoming many frustrations and solving many unexpected problems.

- Even blocks world configurations produce complicated images, confusing programs with mutual illumination effects, indistinct interior edges, misleading scratches, the large variety of vertex configurations, and so on.

- Image input devices like vidicons and image dissectors do not produce the simple clean images one might suppose -- one must understand and account for variations in intensity sensitivity across the image, errors in image coordinate information, electronic noise, light source hum, and lack of dynamic range.

Vision, it was discovered, is very hard.

We Have Scratched The Surface

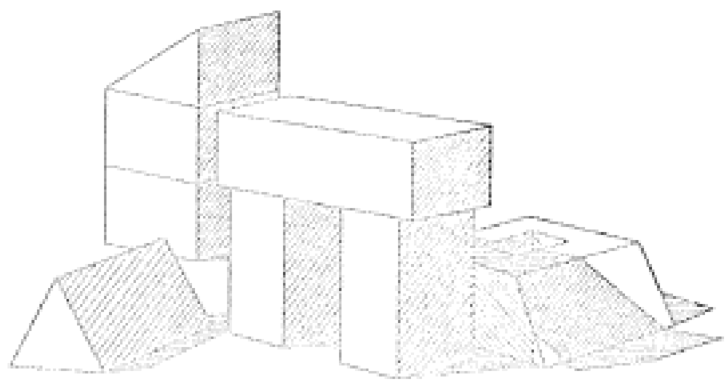
To establish a context for what we propose, we list a few major steps forward in vision research. This is not intended to be a complete listing of achievements over the years -- progress reports are the proper vehicle for that. Instead this is a sketch showing the shift in approach that has resulted from our progress. Vision is no longer naively thought to be a matter of a smart problem solver looking at a few salient image points -- instead it is clear that knowledge representation, constraint exploitation, and care with low level vison are all critical.

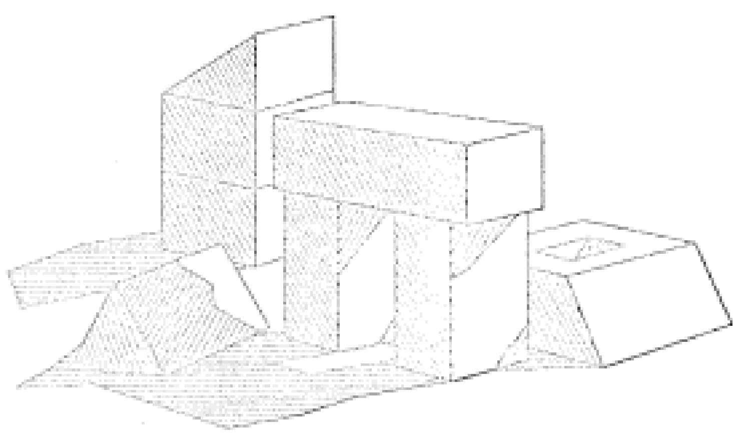
Moreover, it has become clear that the low level visual processing we talk of will lead to techniques that demand a great deal of computational power. Thankfully IC technology is progressing very quickly, so that a real time, flexible visual processor is becoming a technological possibility.

Given the research described in this section, it should be practical to look at photographs of buildings and describe them well in five years. This research into basic issues that confront general-purpose visual processing will complement what we are doing now by way of special purpose vision, which has already begun to produce the prototype level results described elsewhere in this proposal.

We Know That Constraint Exploitation Is The Right Way To Deal With Some Problems Of Analysis

Waltz' work on understanding scenes surprised everyone. Previously it was believed that only a program with a complicated control structure and lots of explicit reasoning power could hope to analyze scenes like that in figure 9. Now we know that understanding the constraints the real world imposes on how boundaries, concave and convex interiors, shadows, and cracks can come together at junctions is enough to make things much simpler. A table which contains a list of the few thousand physically possible ways that line types can come together accompanied by a simple matching program are all that is required. Scene analysis is translated into a problem resembling a jigsaw puzzle or a set of linear equations. No deep problem solving effort is required; it is just a matter of executing a very simple constraint dependent, iterative process that successively throws away incompatible line arrangement combinations. The implications of this type of analysis are far reaching, not only in vision, but also in natural language. It is now apparent that the important task of determining the role of a noun group in a sentence -- agent, instrument, destination, conveyance, or whatever -- can be done by just such a constraint process exploiting properties of word order, prepositions, individual verbs, and global context. The idea is a key element of Marcus' wait-and-see sentence parser, mentioned in *APPLICATIONS ORIENTED BASIC STUDIES* and described in an appendix. Even common sense electronic circuit reasoning has been shown by Sussman to use a similar process. His work is also described in an appendix.





We Know That Understanding The Mathematics Of Illumination Enables The Creation Of Important Analysis Programs

To understand the real world, we must have a different set of primitives from the relatively simple line trackers suitable and sufficient for the blocks world. In the past, image intensities have been used primarily to segment images using *ad hoc* methods based on differences in average image intensities or some higher order measure. This is not good enough. In order to deal with problems in image analysis it is important to understand how images are formed. It is not sufficient to try myriad processing techniques borrowed from other fields in the hope that one can be found that does something useful.

Instead one should realize that images carry a great deal of exploitable information about the three-dimensional nature of the surfaces of the objects imaged. Horn has demonstrated this when he showed how to unwind the differential equations of image illumination in a way that allows shape to be recovered from shading gradients for simple reflectivity functions. Later on, working with Marr, Horn generalized a notion conceived by Land and showed that the natural reflectivities in a jumbled array of colored papers can be determined despite complications from uneven illumination which varies across the image. Such work must be continued. We must develop an understanding of the nature of our visual invironment and the image formation constraints if we are to successfully model what is being seen.

We Know Knowledge Is An Important Adjunct To Image Processing Techniques.

Figure 10 illustrates the action of Shirai's line finder, representative of the work applying problem solving and control techniques to image level vision. Its fundamental thesis was that knowing a lot about the objects in

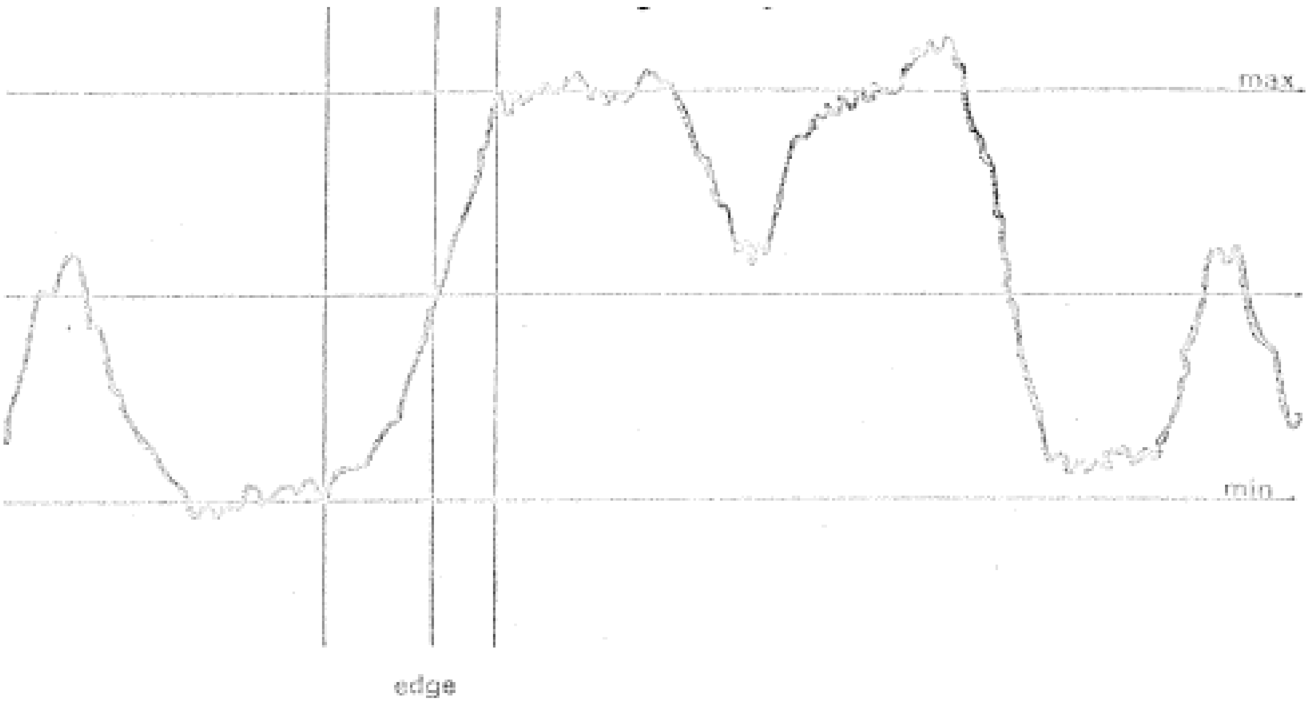


Illustration of the waveform and measurements taken to determine a hypothesized edge position.

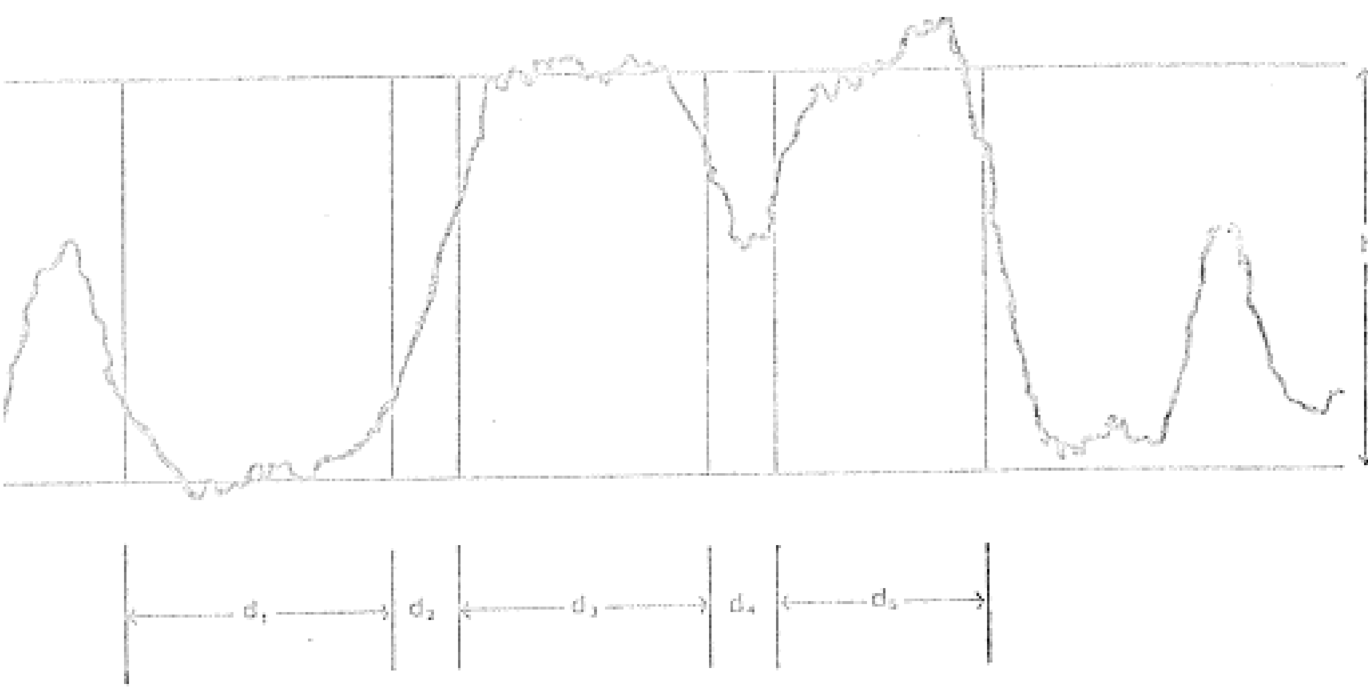


Illustration of the waveform and some of the measurements taken to verify a hypothesized edge position.

a scene makes it possible to analyze the scene with primitives that otherwise would prove inadequate.

Shirai's program knows, for example, that the objects it works with are typically brick shaped and produce many parallel lines. It also knows it must expect aligned objects stacked on one another. Such facts enable it to guess at the location and direction of faint interior lines after it has traced out the strong background lines. Understanding the world should help process real images as well: knowing the approximate place and nature of some clear features like the doors and windows in a room should help locate walls, clocks, pictures, and so on.

Shirai argued that his system illustrates some rules of program organization and style known popularly as "the heterarchial approach." Lately the development of heterarchial philosophy seems to have reached a natural plateau with the work of Freuder. His thesis, submitted in June, 1975, develops heterarchial control structure ideas to new levels by presenting a collection of ideas about how systems can make effective use of suggestions and advice: a hammer head, for example, proposes a possible location for its handle. Such systems have been much talked about and sometimes implemented in simple ways for simple situations, but Freuder's result is a theory rather than an isolated illustration. Freuder develops such ideas to the point where it seems unlikely that further progress can be made in the problem-solving/heterarchial control direction without vastly strengthening the processes that translate image intensities into the symbolic assertions required by any sort of reasoning program.

We believe more work on control structures for vision will not make sense until our understanding of low level symbolic processing catches up. The real world differs from the blocks world significantly because there is a need for much better analysis between the low level feature analyzers and the high level problem solving. We turn to this next.

We Know That The Translation From Intensity Arrays To Symbolic Topography Maps Is Far More Important Than Was Originally Thought

To step out of the blocks world into the real world requires more than problem solving, heterarchy, high level constraint, understanding, and isolated image processing techniques. Past progress in artificial intelligence is a necessary but not sufficient foundation for supporting image understanding systems capable of analyzing and describing pictures. We have come to understand that human quality vision requires talents not at first anticipated. It is these that we propose to work on now.

Our central thesis is that if we are to understand how to make a machine see, we must get our hands the whole way around the problem, not just the so-called "high level" parts of it. We know this in part because efforts to rely on the high level work have not produced satisfactory results, whereas new work on low level vision by Marr and his students, done with a view toward making each level in the analysis as self-contained as possible, has sparked excitement. These results and the proposals that grow out of them form the core of our planned work on basic studies in vision.

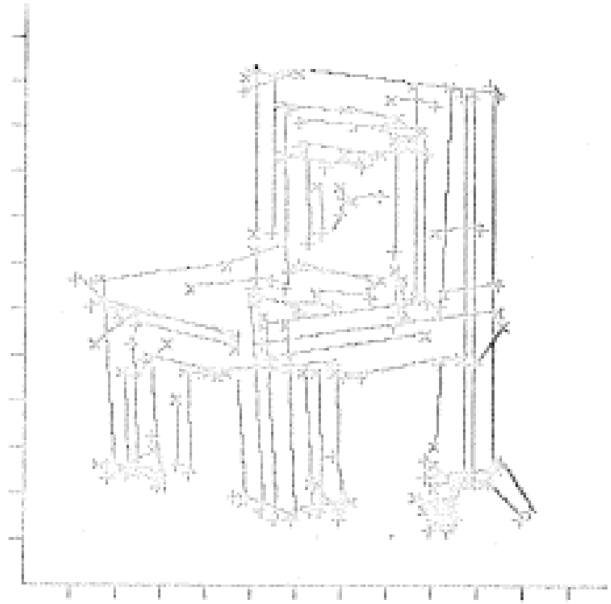
A Solid Theory Of Vision Is Coming

Here, in synopsis, is how Marr feels about vision:

- The function of early visual processing is to compute a description of the gray-level changes present in an image in terms of a vocabulary of gray-level change primitives. These primitives consist of straight contour segments of various kinds (SHADING-EDGE, EXTENDED-EDGE etc.,), LINEs, BLOBs, and of various parameters bound to them such as FUZZINESS, CONTRAST or LIGHTNESS, POSITION, ORIENTATION, simple measures of their *SIZE*, and a specification of their TERMINATION points. This primitive description is obtained from the intensity array by knowledge-free techniques, and it is called the *primal sketch*. The illustration in figure 11 hints at the density of the information taken from the image and represented in the primal sketch. The primal sketch differs from an array of feature points in a subtle way, which is explained fully in Marr's appended paper.

- From our ability to interpret drawings, one may infer the presence in our perceptual equipment of symbolic processes that are capable of grouping lines, points and blobs together in various ways. Non-symbolic techniques, like examining the power spectrum of the spatial Fourier transform of the drawings, cannot account for these grouping phenomena, since the groupings are performed by mechanisms of construction rather than mechanisms of detection.

- For most images, the primal sketch is large and unwieldy. However, it can be capably analyzed by a mechanism that has available the symbolic processes discovered in the previous step, together with the ability to select items out of the primal sketch on the basis of simple first order discriminations acting on the principal parameters. Hence, it is argued, texture vision rests on grouping operations together with first order discriminations operating on the primal sketch, rather than on operations on pairs of points in the intensity array as suggested by Julesz. It is further argued that the set of processes whose existence is necessary in order to explain our ability to interpret a.



b.

COARSE IMAGE DESCRIPTORS
(used in primary control of texture analysis)
Orientation Buckets are 15° wide

ORIENTATION (degrees)	0	15	30	45	60	75	90	105	120	135	150	165
NUMBER OF ITEMS	16	12	2	3	2	1	30	4	2	5	4	14
TOTAL CONTOUR LENGTH	258	264	15	25	14	10	998	34	23	46	25	207

drawings is also sufficient, when applied to the primal sketch, to explain the range of texture vision that is present in humans. Fourier and power-spectrum techniques on their own are certainly deficient, and probably also unnecessary.

Certain fish, incidentally, can mimic their surrounding environment very well with an adaptation time of only a few seconds. Marr's texture theory has led to a theory of the control of this texture camouflage. Presumably it depends on measuring and matching the small number of global texture statistics that Marr believes defines texture space. This theory is being tested elsewhere at M.I.T. under the direction of J. Y. Lettvin.

- The extraction of a form from the primal sketch using these techniques amounts to figure-ground separation. Except in difficult cases, this extraction can proceed successfully without calling upon higher level knowledge, and it precedes the description of the shape of the extracted form.

This has two important consequences. First, the isolation and delivery of a form to subsequent processes does not depend on being able to assign an accurate high-level description to it; and second, because of this it is easy to compute rough descriptions of complex forms. This is probably essential for the fluency of subsequent shape analysis.

- The extent to which higher level knowledge and purpose influences the processing up to this stage is very limited. There is at present no reason to believe that higher level knowledge is needed to compute the primal sketch at all; its role in the extraction of form from the primal sketch often can be limited to deciding which form should be extracted.

It is conjectured that in all cases, higher level knowledge need be only weakly coupled to the processes that separate figure and ground. This relegates the use of higher level knowledge to a much later stage than is found in current machine vision programs, and simultaneously confines its impact to influencing control, rather than interfering with the actual data-processing that is taking place lower down.

Marr's appended progress report explains the theory in more detail.

Progress In Vision Involves Several Goals

- Complete the implementation of the low- and intermediate-levels of the visual theory outlined above. Rewrite grouping code for the sake of speed and assemble the whole, large collection of routines into a single package. The grouping processes developed so far involve an interesting mixture of local analysis, together with an ability to generate one or two more global image variables.

- Devise and implement a theory of lightness computation that works after the extraction of contours. Various motivations exist for this, partly difficulty in doing it at retinal level, partly the feasibility of doing it at this stage provided that (as now seems possible) the image is divided up into regions using knowledge-free techniques.

- Increase the dimensions of the image that we can handle. At present, our images are 128 by 128 elements, which corresponds to viewing an area 1 inch square at a distance of about five feet with the human retina. Even this small image stretches our resources, and processing one such image can take 20 or 30 minutes of CPU time. If we spend considerable effort on re-writing chunks of the low-level code, we could probably move to a 256 by 256 image with only a small added cost in processing time.

- Put BLOBs securely into the primal sketch. A BLOB is a small, indivisible patch whose lightness differs from that of the surround. Although it is clear that BLOBs should be a primitive element, the existing techniques for detecting them from the output of the orientation-dependent analysis are somewhat elementary.

- Devise a theory of the alignment of an axis to an extracted form. This is vital for the ideas we are currently developing about the transition from two to three-dimensional descriptions. One subgoal here is to devise and implement a theory of symmetry detection because this is one way of providing an axis.

- Carry out experiments to discover whether the theory provides a complete account of texture discrimination abilities. This is a difficult and arduous undertaking since we are really searching for a counter-example. But analysis of (say) a hundred images, chosen for their awkwardness, would go a long way towards establishing adequacy. At present, about 20 have been processed, and not all our texture-processing facilities have been used.

- Assemble a comprehensive report of the complete low-level visual processor. It would initially be published as a Technical Report, then a published monograph.

We wish to express in the strongest possible terms our view that it is unrealistic to expect useful results from most components of a photo-interpretation project until at least this level of lower-level technological competence has been achieved. We cannot help but feel that those who believe all parts of this immensely difficult problem can be solved using existing techniques and machines have an insufficient understanding of the magnitude of the issues that are involved. On the other hand, we do feel that a careful and energetic basic research approach to the vision problem, in which one addresses the issues that arise as they become accessible for solution, is the correct long-term scientific approach. Problems will not be solved just because one wants to solve them; they also have to be ready for solution. If we pursue the fundamental issues with diligence and vigor, then we may reasonably expect that within a period of about five years, some of these very advanced problems will indeed come into development, rather than speculation range. But bear in mind that this cautious statement is in fact extremely radical, and would have to have been regarded as wildly over-optimistic as little as one year ago.

Most Of Our Stated Objectives Will Require Two Years

Barring unforeseen delays, we should be able to complete much of this proposal within two years. Many of these problems now require labor more than ideas, and unless we have overlooked something important (which is possible), we should have completed the low- and intermediate-level studies by the end of 1977.

There are, incidentally, some problems we would like to solve but cannot approach yet. These problems include those of understanding the visual qualities of gloss, glitter, shinyness, wetness, metal, reflections, brilliance, transparency, and translucence. We feel that many of these attributes in an image can be detected using simple autonomous techniques, but as yet our grounds for this are speculative, resting mainly on the intuition that they generate recognizably distinct "qualifiers" in a perception, rather like color does. Ullman's study of fluorescence strengthens this feeling because he showed that an autonomous, knowledge-free method can detect light sources.

Thinking About A Vision Machine Is A General Medium-Term Goal

Once a system has been implemented in a single, autonomous package, we shall be able to study higher-level (3-D) issues in the context of what such a system can deliver from raw image data. The sheer computing power that is needed to carry out texture processing on an image of reasonable size means that we should begin thinking about designing some special-purpose hardware that is powerful enough to carry out the processing in a few seconds. Accordingly, we have begun preliminary design studies on such a machine, and we wish to continue these. The two prerequisites for serious proposals for construction are (i) a firm demonstration that the PDP-10 implementation of the whole processor performs satisfactorily; and (ii) that our design studies result in a solution to the hardware problems that lies within what is technologically possible, and which will perform the whole operation within 20-30 seconds; and (iii) that the expense is reasonable and nothing like the multimillion-dollar parallel machines of the past. Unless unforeseen difficulties arise, we expect to have a concrete design proposal, together with proof of its processing adequacy and computational prowess, by early 1977.

COMMON SENSE REASONING

It is not possible to feel comfortable with the problem of common sense reasoning. Progress has been made by Sussman, Goldstein, and others like McCarthy at Stanford and Newell at Carnegie-Mellon, but some very basic work remains to be done.

Since Slagle's integration program helped establish artificial intelligence as a serious science, researchers have mined out more and more narrow domains by ferreting out what kind of knowledge is involved and how to use it. A correct complaint rests on the observation that one program contributes to the next only on the level of ideas and methodology. Something better is needed. Collecting a group of specialist programs together for a particular kind of problem is good and helpful but not a full solution.

Thinking about common sense reasoning seems to be the right thing to do. This is not a new theme, nor is lack of a better theory an obstacle to the first round of progress on the Very Large Data Base problem or the Intelligent Terminal Problem. On the other hand, we believe that without basic research in this area, progress on practical problems would stop short of a full solution. Systems would tend to be constrained to the specific chosen implementation domains with transfer to other areas remaining difficult and time consuming. A system worked out for medicine would be useful to the creation of a system for ship maintenance, but much new work on the epistemology of ship maintenance would have to be force fed.

But what is the alternative? *We believe that understanding common sense reasoning is necessary before systems can be created which learn easily about specialized problem solving domains.* Trying to work out mechanisms that draw direct analogies between heart disease and the maintenance of boilers might be one approach. But to us, a better idea is to discover how both might be understood in terms of the epistemology of the simple, everyday physical world. *We therefore believe a key need is a theory of common sense reasoning about simple physical mechanisms.*

Sussman and Goldstein have made a good start by working out notions dealing with process: we now have a better vocabulary for talking about goals, plans, and the structure of programs. What we need now is a complementary effort focused on our physical environment. We need to have a better vocabulary for talking about objects and materials. We need to apply our classical methodology to the simple physical world interaction problem. We need to establish what kind of knowledge is needed, what specifically is needed, and how much. We need to work out the epistemology of interacting things -- things pushing one another, the idea of filling, filling to store and filling to inflate and filling to the bursting point as in hearts and boilers. Our systems need to know about the ideas of repair, patch, and breakdown, the toughness of common materials, the idea of floating, rolling, mating, helping.

The observations of Martin at MIT's Project MAC and of Schank at Yale reinforce the idea that specialized reasoning is enabled or at least facilitated by analogy to simple physical happenings. Everyday language is replete with physical world terms and usage applied to abstract facts and processes. In the mental world we talk of ferreting out facts; in Martin's social world ownership moves from place to place with language suited equally to objects; in the logical world facts support one another in a way not unlike Winograd's blocks; and for a typical expert problem solving domain, it is rare to think of learning electronics as a child without thinking about resistors in terms of narrow water pipes and capacitors in terms of water balloons or some such reservoirs.

Given such a world theory we believe we will be in much better shape to handle all sorts of diverse fields. At first the system may not conclude for itself that a heart can be understood in the same terms used to understand the atomic vessel and fluid mechanism, but if we can tell the system such things, much needless reinstruction and narrowly focused coding would be circumvented.

There Are A Number Of Good Precedents For Developing A Theory Of Common Sense Reasoning

- The Ph.D. theses of Sussman and Goldstein probed the epistemology of processes, showing that considerable regularity emerges when the right descriptive apparatus is available. Sussman worked with the automatic debugging of blocks world motion primitives *a la* Winograd, and Goldstein handled simple geometric figure programs.

- Sussman, Brown, and McDermott's work on electronics carries this thread forward and gives us considerable understanding of procedural aspects of the problem and of the interfaces between complex parts.

- Schank's conceptual dependency theory, although a bit overfocused on the human body, is also an intellectual precursor. The important feature is the demonstration that convincing broad-based formalisms are achievable. We expect our theory to be considerably more complicated, but not impossibly so.

- Papert uses the term "qualitative physics" to cover the problem of understanding the pendulum and the gyroscope in simpler, push-and-pull terms. In his NSF supported work in education, he has worked out ways of looking at these things that will be helpful.

We Plan To Attack The Common Sense Reasoning Problem From Four Directions

In summary, the long range objective of our work is an understanding of the simple physical interactions of simple physical things sufficiently good to serve as an analogical foundation for programs that want to reason about more abstract problems. (For full power, the analogy process eventually may want to be capable of working between two specialized domains as well as back to the fundamental physical one.)

We believe this is deep research and some breakthrough requirements prevent us from giving honest estimates of advanced performance levels and dates. We do believe, however, that there are good ideas about the next steps and that work on these in the two year time frame will leave us ready to make the predictions that would be specious now.

In particular, we feel that the next steps must be directed to work on these four subjects:

1 Freiling wants to understand the primitives that are prerequisite to understanding basic machinery. He talks in terms of shafts, surfaces, wedges, and a dozen or so other similar primitives. He further mentions interactions like rolling, sliding, and the like. And putting everything together he wants a system that can understand how a camera works by drawing analogy between light and fluid flow, realizing in consequence the need for a container for the film, a conduit for the light, and a part to control its "flow." This is ambitious, and two years' work may not produce such an advanced scenario, but the direction seems correct and some analogical reasoning ability seems certain. Freiling's notes on approach are appended.

2 DeKleer's study of qualitative reasoning complements Freiling's by attacking the question of how to deal with the boundaries that segment complicated problems into simple pieces within which straightforward mathematical treatment suffices. Reasoning of the sort DeKleer is getting at pervade our thinking whenever we must have more than a rough idea about how a complicated machine works -- surely it is a prerequisite to design and must be worked out if machines are to be designed automatically.

3 Richard Brown probes the general question from a third direction, that of concentrating exclusively on using analogy in reasoning. His approach is just now taking shape and now it can be said only that solid geometry may be the domain because a great deal is known about solving tough solid geometry problems by using simplified plain geometry problems together with analogy. This domain will be pursued only if it remains promising as a domain in which to study the general analogy problems. The geometry, *per se*, is of secondary interest.

4 Kuipers is working on representing the kind of knowledge that people have in their mental maps of cities. This is the knowledge which is used to represent where things are with respect to each other, how to get from one place to another, and what kinds of groupings they fall into. His research will shed considerable light on the issues of imagery, visual memory, visual problem solving, and similar issues involving the internal representation of spatial knowledge. There is considerable psychological evidence that for humans the mechanisms used for representing and manipulating such knowldedge is extensively reused when dealing with more abstract topics like the world of organizations. Kuiper's project complements not only the program aimed at dealing with common sense reasoning but also the work of Lavin which is oriented toward extracting information from scenes with hills and lakes. Lavin's proposals are described in *APPLICATIONS ORIENTED VISION AND MANIPULATION*.

EXPERT PROBLEM SOLVING, LEARNING, AND THE EPISTEMOLOGY OF ENGINEERING

A long range goal of Artificial Intelligence research is to make a system which can assimilate new knowledge and learn from experience. In reaching for it we have seen many attempts at producing universal formalisms. Each of these has made the fundamental assumption that knowledge is additive; that means can be found for assimilating new knowledge without overhaul of the existing knowledge structure. People who design problem solvers are not alone in making this fatal assumption; many students believe that they can learn by just listening to facts thrown at them in lecture halls and that learning can happen without much intellectual effort. The truth is that learning involves great effort. Introduction of a new fact or concept into a knowledge structure often precipitates upheavals and dislocations in that structure. Any system which can assimilate new facts must be competent to modify and manipulate its existing knowledge.

Engineers do this constantly, they combine, analyze, debug, and explain stuctures in the course of design. They decide how simpler structures can be combined to achieve particular goals and they predict the behavior of composite structures by combining the more primitive behavior of the substructures out of which they are formed. Then they use analysis because proposal of plausible designs may lead to designs which do not quite work. This is discovered by analysis and leads into the debugging phase of design. Finally, an engineer must be able to explain the devices which he has designed. An explanation is often a description of how the behavior of the composite device can be attributed to the combined behaviors of its parts.

We Want To Understand Programming In Terms Of Engineering Design

Since some knowledge is inherently procedural, and thus is best represented as programs, we are forced to the conclusion that a learning program must be a programmer -- it must understand knowledge represented as procedures, so as to be able to modify and patch them to incorporate desired changes.

- Consequently we must have a sound theory of programs and how their structure relates to their functions as defined by the goals of their designer. We are led to a study of the epistemology of the design of programs.

Can we write programs that combine, analyze, debug, and explain other programs? We have begun to, but we do not know how to properly describe a program and its relationship to its design. Programs are hard to document in clear and useful ways. In order to write programs that reason about programs we must somehow learn to describe programs in terms of the designer's goals and plans and the interactions of the parts. The research we propose to do is aimed at clarifying these concepts.

We believe that some general programming knowledge is composed of "plan fragments" which are abstract and rather general plans for accomplishing goals. An example from Sussman's blocks world program writing system is: "To achieve x and y, try to achieve them separately." Another simple example is the fragment: "To achieve x, if we have a method of achieving x given that we have y, first achieve y, then use the method for x." Still another is: "To achieve 'There are no x's with property P,' find each x with property P and achieve the negation of P on x." These rules are in fact only heuristic and likely to result in bugs. Though the fragments shown are abstract, fragments are intended to represent concrete knowledge as well. Programs are constructed by expanding and instantiating plan fragments. At each level, which of the applicable plan fragments is used to achieve a particular goal is a design decision. A plan for a program is, at least in part, the tree of expansions of plan fragments, with particular design decisions, used to create the program. Since plan fragments often expand into bugs, we also need notations showing what constraints were added (such as ordering) which were not parts of the original design, and what bugs they were supposed to eliminate.

Programs Are One Example Of The Notion Of Deliberate Artifact

It is probably clear that plans for programs can become very complicated and studying the epistemology of programs directly may be too hard. Programs have conditional control structures -- their configurations (how they break up into parts and how these parts are related) may depend strongly on their inputs. They have shared structures such as subroutines and a modification to such a substructure potentially affects all of the places that substructure is used. Programs also deal with rather complex encodings of information -- there are many kinds of data structure a programmer must know how to manipulate.

These facts are discouraging, but it has become clear to us that the notions of plan fragments and plans, and using such concepts in design, analysis, debugging, and explanation, are definitely not unique to the domain of programming but are common to reasoning about any "deliberate artifact." We have come to feel that we can make substantial progress in the study of programs and their properties by first examining the nature of these concepts in other, simpler domains.

There Is A Continuum Of Domains For Studying The Epistemology Of Deliberate Artifacts

We have searched for a simple domain which eliminates many of the problems posed by the study of programming but which also deals with complex deliberate artifacts composed of simpler artifacts deliberately arranged. In our search we wanted to avoid several difficulties: the problems of spatial reasoning, visual recognition, and natural linguistics, all of which involve other deep, deflecting issues.

As we see it, all of this argues for diving into the following at a carefully selected point:

- The simple and much studied blocks world.

- One of the many engineering areas that can be approached through the relatively simple mathematics of lumped parameter linear systems. These include the design of electronic circuits, many simple mechanical devices, and hydrolic systems.

- One of the many engineering areas that makes use of interconnections of subsystems characterized as ported, linear, black box information or materials processors. The lumped parameter linear systems mentioned above are to be understood to be a subset of this class for which the port descriptions are constrained to a certain simple form. Almost all forms of engineering have a strong involvement with this kind of analysis in the course of design.

- One of the areas in which the black boxes are non-linear.

- Programs.

In the work done so far, the concentrations has been on the design of lumped parameter linear systems for several reasons. It is well understood by engineers of all sorts and it has a precise technical vocabulary along with interesting problems. The criteria for success are clear. The lumped parameter elements interact in stereotyped ways only at designated ports. Even meechanical engineers, who are often stuck with geometry, still make use of these principals although they are limited in the complexity they can successfully handle relative to, say, the electronics circuit designer.

Much Is Needed

Our work to date, executed in the study world of electronics design, has confirmed a need for the following obvious things:

- The need for knowledge representations adequate to define system requirements.

- The need for transformations among representations, in this case magnitude graphs to pole-zero plots to algebraic system function form, for example.

- The need to match plan fragments in the bag of tricks against system requirements. There is an implied need to understand how to match when the matches are not exact.

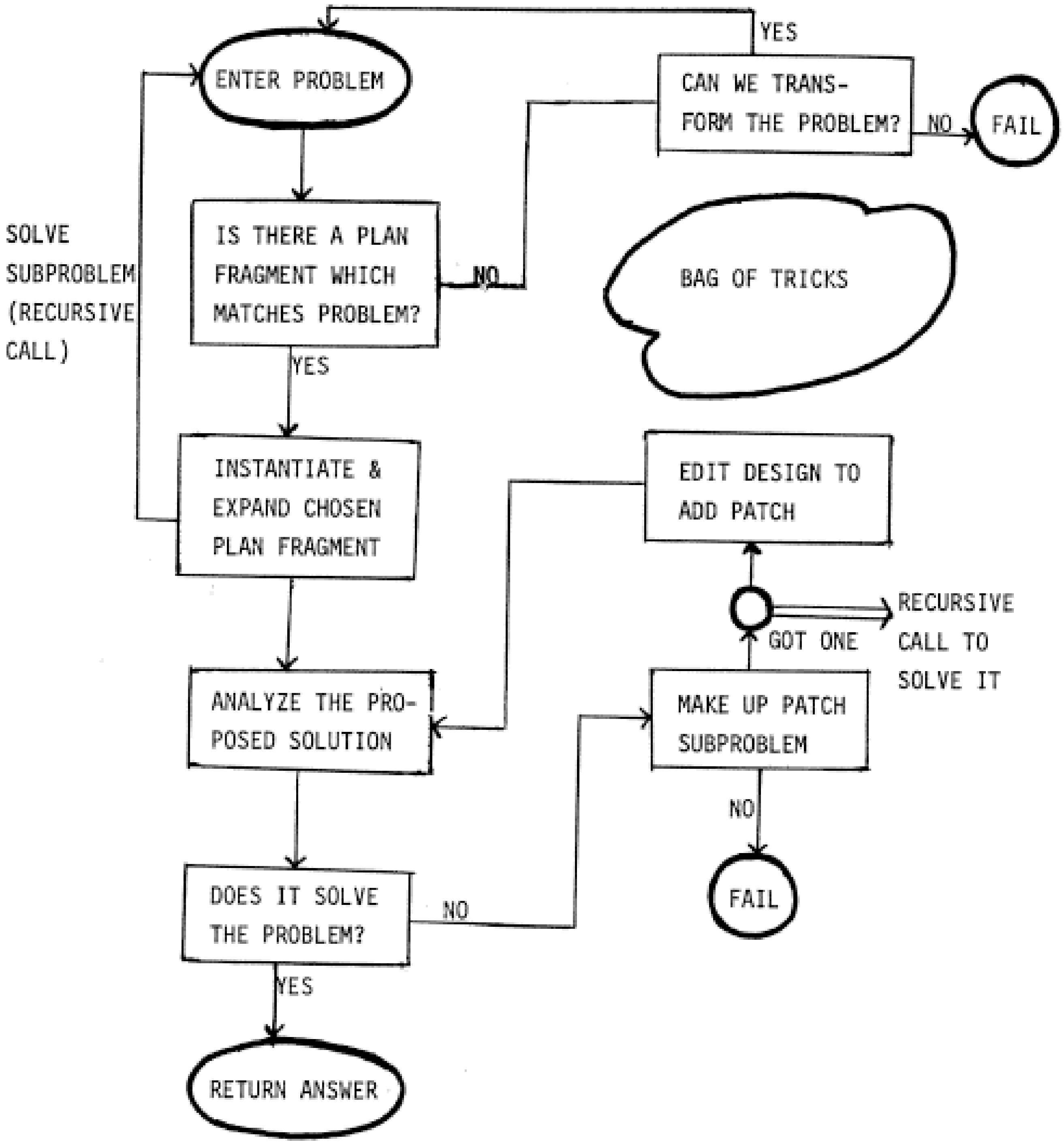
- The need to debug the undesirable interactions that follow from first approximations which ignore fragment interactions. In this case, the loading of one stage on another.

All these are necessary, we believe, to expert problem solving in general since they have analogues in any domain that could be classified as a domain of deliberate artifacts.

Figure 12 shows the kind of control structure we feel is required for the kinds of problems we have in mind. Note that there is often a "Bag of Tricks" -- a data base referred to by all parts of the program.

There Is A Plan For Understanding Deliberate Artifacts

The Engineering Problem Solving Project is concerned in the short run with electronic design (either IC or discrete). Later it is to blend into an attack on the structure of more complicated artifacts, such as programs. In this section, we outline our timetable.


ENTER PROBLEM
IS THERE A PLAN FRAGMENT WHICH MATCHES PROBLEM?
NO
YES
CAN WE TRANSFORM THE PROBLEM?
YES
NO
FAIL
SOLVE SUBPROBLEM (RECURSIVE CALL)
BAG OF TRICKS
INSTANTIATE & EXPAND CHOSEN PLAN FRAGMENT
EDIT DESIGN TO ADD PATCH
RECURSIVE CALL TO SOLVE IT
GOT ONE
ANALYZE THE PROPOSED SOLUTION
MAKE UP PATCH SUBPROBLEM
NO
FAIL
DOES IT SOLVE THE PROBLEM?
NO
YES
RETURN ANSWER

- Gerald Sussman, Richard Stallman, and Gerald Roylance have written and extended EL, a program for circuit analysis. This program uses its knowledge of the plan structure of the analyzed device to avoid blind relaxation techniques in analyzing circuits. As of now, EL handles quiescent circuit analysis. Sussman *et al* are currently engaged in extending EL to AC signal analysis and reorganizing it toward greater flexibility. By June 1976 they plan to have a finished product. EL's ability to "explain" its conclusions will be critical. The system will not only give answers on demand, it will explain those answers as well.

- Currently, good progress is being made on electronics design debugging. Allen Brown has finished a Ph.D. dissertation on "Localization of Failures in Radio Circuits." This work will be extended by Matt Mason, who will implement Brown's ideas. He should have Brown's LOCAL program coded by June, 1976.

- Drew McDermott is studying the problem of interface design. He has written a very simple program that knows about a few of the linear system plan fragments. There are two main subproblems that have arisen: choosing transformation techniques for problems; and satisfying conflicting constraints in putting plan fragments together. He will complete a Ph.D. dissertation on these topics by June 1976. McDermott is doing his research under a commitment to a certain kind of modularity in programming. He would like his program to have "wired in" the theory of artifact design; but the plan fragments, problem analysis techniques, and other "advice" are to be more loosely coupled to the program.

- There remains the goal of understanding the design and modification of complex programs. In parallel with our electronics work, Charles Rich and Howard Shrobe have been studying computer programs directly, in terms of engineered artifacts, like circuits. They are aiming at a program to aid a human in designing hash tables, by June 1976.

- The Rich and Shrobe work is part of a larger goal of the development of design analysis tools for programs analogous to the analysis tools used in lumped parameter and linear systems analysis.

For reasons we have mentioned, programming is a very difficult domain. It is hard now to predict what we will have learned from another engineering discipline. In a year, we will be in a good position to bring our intuitions and experiences to bear on program design.

We believe that the work outlined in this section is one component of what is necessary before we can design systems that can really write their own programs. The other component needed, it seems to us, is the thorough understanding of the domains for which the automatic systems are to programmed. This is one reason why we applaud Martin's work in Project MAC. We believe our effort to be complementary to his.

PERSONNEL

- Work in basic vision studies is under the general direction of DAVID MARR. His theories of vision are on the critical path to a better understanding of vision. Professor BERTHOLD K. P. HORN, whose first responsibility is to direct the applications of vision and manipulation work, will continue to contribute from time to time.

- SHIMON ULLMAN, who just recently completed a sucessful study of light source detection, will continue to pursue aspects of the Marr theory. Other students with likely involvement include KEN FORBUS, an undergraduate who has helped Marr with programming for some time and JAN GALKOWSKI, a graduate student who has just completed a study of electronic component identification.

- Professor PATRICK H. WINSTON plans to take an active role in working with students on the common sense physical reasoning problem.

- MIKE FREILING has proposed to work on the common sense reasoning and analogy mechanisms appropriate to basic machinery. RICHARD BROWN, hoping to deal directly with the problem of analogical reasoning, will try to expose key issues by studying the way solid geometry problems are solved throgh plane geometry analogies. JOHAN DEKLEER, working under the supervision of Professor SEYMOUR PAPERT, is working on the qualitative knowledge/quantitative knowledge interface. BEN KUIPERS is doing mental maps.

- Professor GERALD SUSSMAN is in charge of making progress in the general domain of expert problem solving, program understanding systems, and the epistemology of "deliberate artifacts." Some of this work is done in cooperation with Professor CARL HEWITT, supported mostly through Project MAC and partly through the Artificial Intelligence Laboratory.

- Sussman and Hewitt are jointly supervising CHARLES RICH and HOWIE SCHROBE, who are working on the problem of expert problem solving using an expert in programming as the discovery domain.

- Sussman has worked closely with RICHARD STALLMAN and GERALD ROYLANCE on the development of the electronics expert, EL. Roylance plans considerable further work in this direction, expecting to introduce small signal analysis to complement the quiescent analysis expert implemented so far.

APPENDICES

MARR: ON TEXTURE AND A NEW THEORY OF VISION

FREILING: ON SIMPLE MACHINES AND ANALOGIES AMONG THEM

DEKLEER: QUALITATIVE AND QUANTITATIVE KNOWLEDGE IN CLASSICAL MECHANICS

SUSSMAN AND STALLMAN: HEURISTIC TECHNIQUES IN CIRCUIT ANALYSIS

SYSTEMS
AND
SYSTEMS
CONCEPTS

SYSTEMS AND SYSTEMS CONCEPTS

The Artificial Intelligence Laboratory wishes to develop a low cost system intended to serve single users with an outstanding LISP together with equally high quality text editing, debugging, and display features. The overall goal is to produce a computing system dramatically superior to any available today both in terms of power and cost. We intend to achieve this goal by way of implementing a prototype system manifesting ideas now worked out and detailed in appendices.

We also intend to do some thinking about vision processors with a view toward preparing for the time when such machines are both necessary and possible.

PURPOSE

This section is short since other appended documents fully convey the general spirit and technical detail behind this effort. These few words will simply summarize the salient points.

The LISP Machine Promises A New Way Of Thinking About Computing

The trend has been toward cheap hardware and expensive software. In contrast to the classical shared CPU and central memory systems, the LISP machine concept exploits the cheap hardware trend. A principal result of the overall approach is simpler system software. The impact on how computing is done will be, we believe, at least as dramatic as the introduction of time-sharing. It is perhaps not surprising that such a development should spring up from the community of people who think about Artificial Intelligence inasmuch as Artificial Intelligence development has always made extreme, technology-pull demands on computer system science. It is argued, after all, that time-sharing itself was largely incubated and promoted by early work in artificial intelligence.

The LISP Machine Offers More Power

The family of ideas that make the LISP machine an exciting prospect anticipate the future. Paramount among these ideas may be the large address space offered. We believe that 1024K of address space will be needed eventually in situations where natural English input is demanded. We base this on estimates that seem sensible in terms of Winograd's natural language work in the simple blocks world which required 150K and on subsequent work by Martin which leads him to talk in terms of 512K as a minimum. Existing and proposed PDP-10 type machines allow only 18 of the 20 bits of address needed, while the LISP machine is to provide 23, a comfortable margin. (The KL-10 processor will address more than 18 bits of address space, but only by way of an awkward afterthought method forced by design decisions which were reasonable at the time the PDP-10 architecture was introduced years ago.)

We believe that natural language is not alone in suggesting the need. So far programs have worked with sophisticated programs on small data bases. As we become more ambitious and more concerned with real problems, the larger address spaces will be not just desired, but essential. We note, as an example of this, that the CONNIVER interpreter alone occupied over 100K of central memory and therefore points toward systems containing it which must be beyond the reach of the 256K, 18 bit address space.

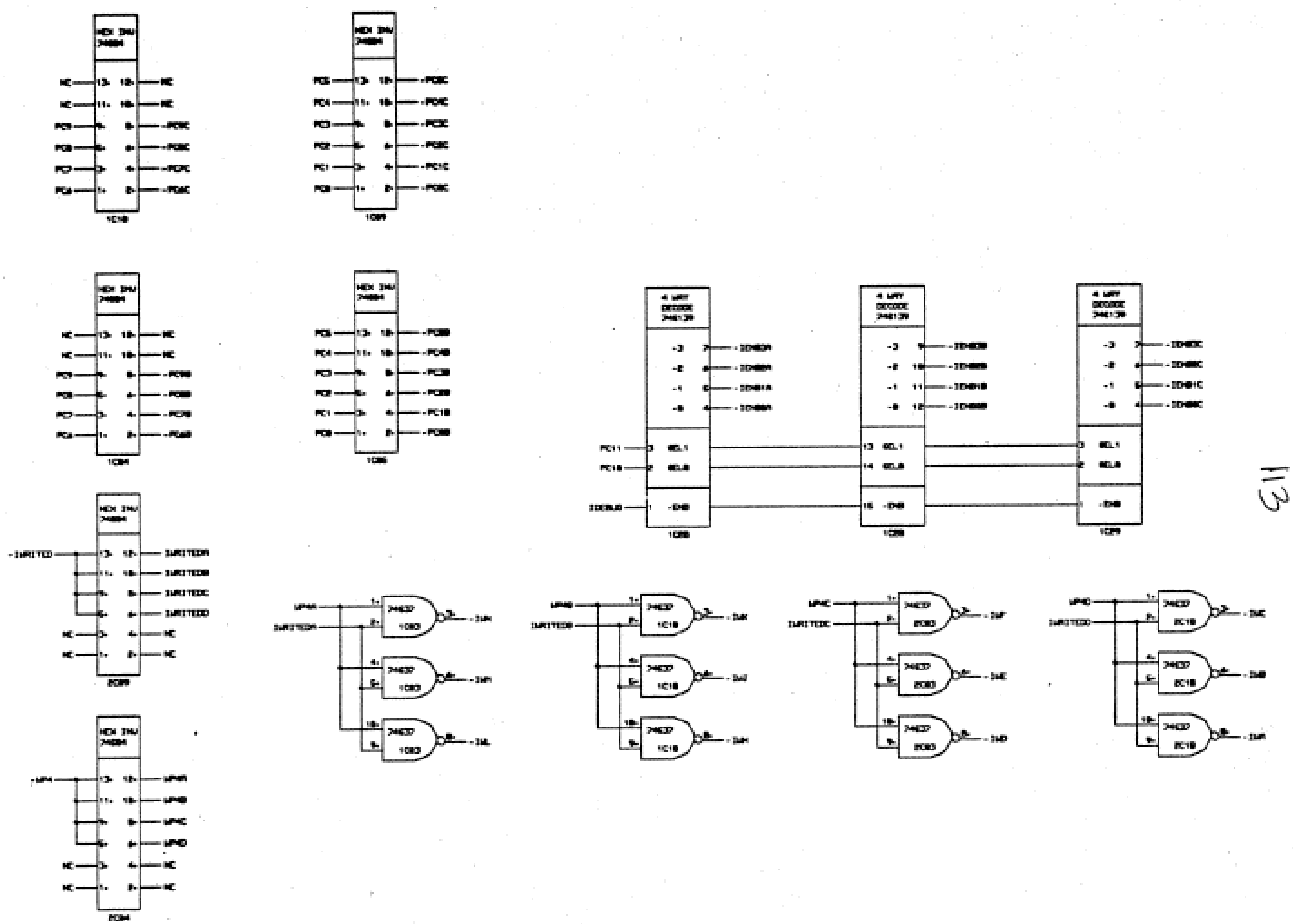
LISP Machine Development Will Be Surprisingly Inexpensive

Ten years ago it would have been foolhardy for a university group to propose the creation of a grand new computing engine. Times have changed. Large scale integration has given us prefabricated walls where before there was only brick and mortar. Five years ago at least ten times as many chips would have been needed. Moreover, Stanford and M.I.T. work on interactive, graphically oriented design aids has given us automatically generated documentation and wire wrapping programs where before we had only drafting mistakes, lost bits of paper, and chaos. Figure 13 illustrates the sort of documentary output automatically generated in the course of the design process. These new tools have been thoroughly demonstrated through use on projects of demanding scope, including the design of the new DEC KL-10 processor.

The LISP Machine Will Serve The Needs Of Many

The time has come to produce such a machine because the technology is ready and the users breathless. Certainly the artificial intelligence R&D community is the first source of customers. For this group, not only are the technical advantages extreme, but also the new ability to get started with an initial investment one-tenth the size of a PDP-10 configuration is extremely attractive. Smooth expansion via one-by-one addition of LISP machines substitutes for big ticket decisions. More people, widely supported, will do Artificial Intelligence work. Of course the scope of the predicted impact reaches far beyond Artificial Intelligence for two reasons: one is that the predicted infiltration of Artificial Intelligence into most of the frontier problem areas of computer science has occurred; another is the fact that the LISP machine itself, while worked out with LISP in mind, will certainly support other languages with largely the same advantages. The concept will make life easier and more cost effective not only for the Artificial Intelligence LISP using community, but also for those who for historical or genuine technical reasons must stick with the more classical languages for production and execution of end user software.

In particular, we believe LISP machines, being very cheap in 10 years



time, will be the hardware foundation for the computer user's INTELLIGENT TERMINAL, the manager's MANAGEMENT ASSISTANT, the student or trainee's PERSONAL PEDAGOGUE, the factory's INTELLIGENT ASSEMBLER, and, in general, everyone's PERSONAL HELPER.

THE LISP MACHINE CONCEPT

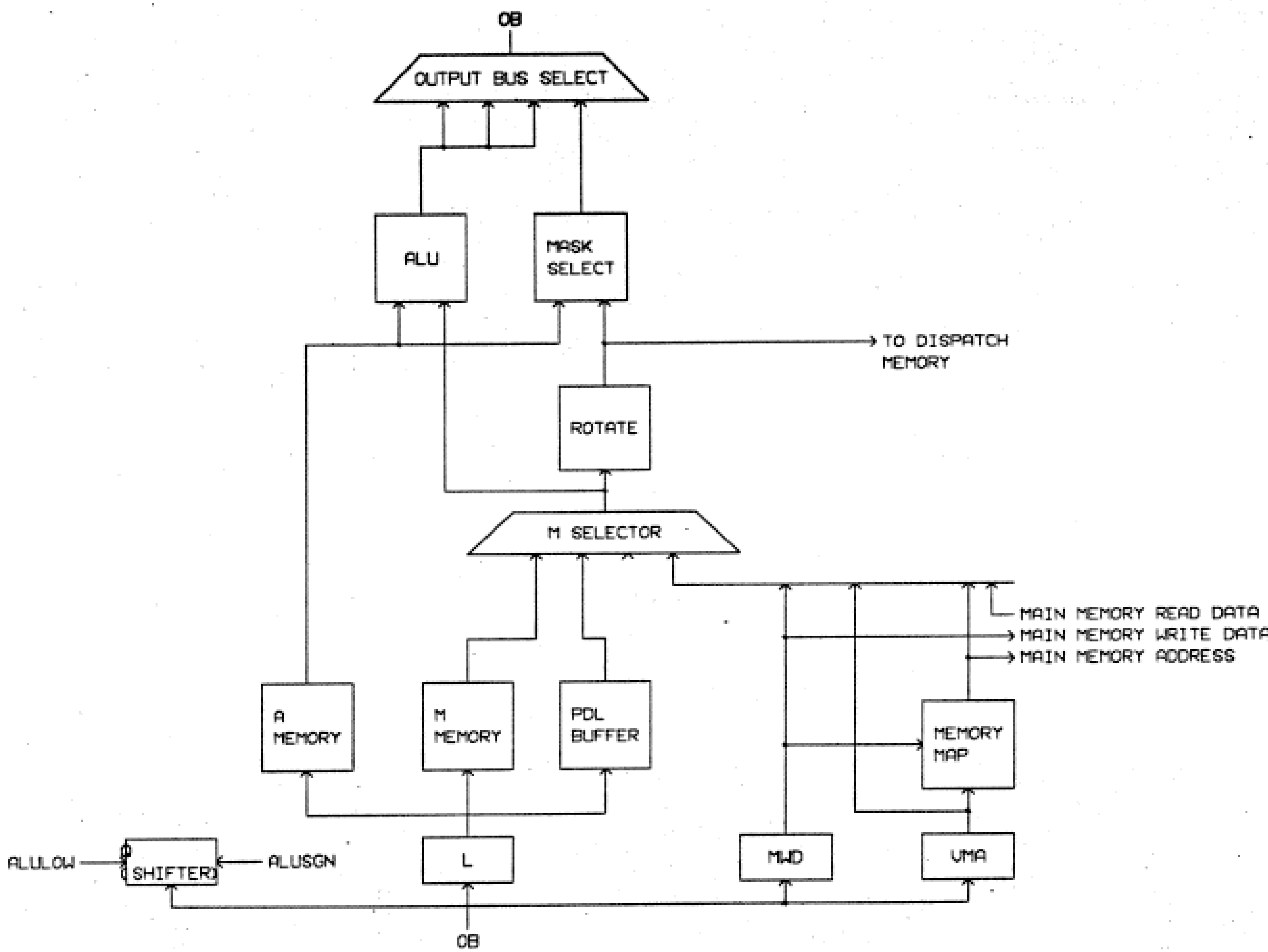
The hardware configuration is diagrammed in figure 14. It is to consist of the following core:

- A PDP-11 for doing I/O, for controlling a unibus, and for aiding in debugging other components, principally the CONS processor.

- The CONS processor for executing LISP tailored instructions. The CONS processor successfully integrates many good ideas, among them programmable microcode, LISP oriented byte handling operations, special list processing directed addressing modes, a fast internal stack, and more. The CONS machine is fully described in an appendix.

- A good deal of central memory, probably on the order of 50K to 100K 16 bit words. This large amount of non-shared, dedicated local memory is determined by the overall design philosophy which rejects classical time sharing in favor of system simplicity and performance.

115



CONS; DATBOX

- A fast local disk or possibly CCD bulk memory of the sort recently heralded by the introduction of 16K CCD chips. This will hold that part of the 23 bit address space not in "fast" memory and will be the primary source of system software and user programs.

- A connection to the rest of the computing world seen by the machine. Here is where the sharing horizon is to stand. Some link, probably some sort of Ether net, connects users together, provides for data and program sharing, and all the usual community features introduced by time sharing.

COMPUTING ALTERNATIVES

Several other competing recommendations for meeting computation needs have been considered. There are good arguments for some although we believe the weight of the evidence favors the LISP machine. Here are the principal alternatives.

- Option 1: Simply wait for the $50,000 PDP-10. The extensive software work already done makes this a reasonable position. Unfortunately the once state-of-the-art PDP-10 architecture now creaks with age. The small address space is particularly worrisome. We do see, however, an intermediate time during which LISP machines will stand parasitic on PDP-10 systems waiting for users ready to run LISP. The PDP-10 will continue to run the excellent text editing and general support software already existing, while LISP execution takes advantage of the LISP machine's dramatically better ideas for running LISP. This is the reverse of the usual idea which farms out the secretarial work to PDP-11s, presumably running yet unwritten secretarial software.

- Option 2: Develop a PDP-11 LISP. We have done this, but intend no major work dependent on it. The address space available to the PDP-11 architecture forbids it. Such a LISP can run some small demonstration and application programs, but only a small portion of those being written now by workers in Artificial Intelligence.

- Option 3: Finish the LISP machine. This seems the wise course.

LISP MACHINE OBJECTIVES

Our intention is that the LISP machine development be completed during the next two years with a prototype running by January, 1976 and an advanced prototype running by January, 1977.

LISP MACHINE HISTORY

Richard Greenblatt conceived the idea of the LISP machine. He was soon working in close cooperation with Tom Knight.

The work proposed here was earlier proposed as a joint project between the Artificial Intelligence Laboratory and Project MAC. Joel Moses of Project MAC wrote the original proposal. As now organized, however, the work will be handled within the Artificial Intelligence Laboratory with a view toward close supporting counsel and communication with Project MAC, which is to be regarded as equally responsible for getting the ideas off the ground. It seems clear that MACSYMA and other Project MAC interests will be best served in the future by the LISP Machine or something very similar.

PERSONNEL

This project is in the hands of people well known for their ideas and preeminence in system software and hardware design. RICHARD GREENBLATT, principal author of the ITS time-sharing system, is in charge of the software. THOMAS KNIGHT, principal author of the Knight terminal display system, innumerable interfaces and special purpose processors, and overall coordinator of our digital hardware systems, is in charge of the hardware. They will work with JACK HOLLOWAY, well known for his work on the so-called Superfoonly Project at Stanford which was extremely influential in the development of the latest in the PDP-10 family of computers, the KL-10, and with JOHN MOUSSOURIS, who is now finishing up a parallel processor for chess (not funded by ARPA). BILL FREEMAN will interface the LISP Machine to the existing time sharing system.

A VISION MACHINE

Soon we will know enough and prices of processors will be sufficiently low for sensible proposals to be made for very general parallel image processing configurations. We are therefore giving some thought to what could be done, particularly with the recently announced LSI-11s. This thinking is not far along and it is inappropriate to make a sharp proposal at this time. We do append a working paper on the subject, however, to show that such a project deserves some thought.

APPENDICES

HORN AND WINSTON: PERSONAL COMPUTERS (REPRINTED FROM DATAMATION MAGAZINE)

MOSES: ORIGINAL LISP MACHINE PROPOSAL

KNIGHT: THE CONS MICRO PROCESSOR

MARR AND FORBUS: VISION MACHINE -- PRELIMINARY ASSESSMENT