March, 1971

DRAFT DRAFT

VISION FLASH #9

WHAT'S WHAT

by

Patrick H. Winston

An outline of the modules used in the copy demonstration, the reasons for doing robotics, and some possible directions for further work.

What has been done and why

There exists a system occasionally capable of more or less routine copy efforts. The system is modular and incrementally expandable. We anticipate no difficulty in generalizing its ability beyond the universe of bricks currently dealt with.

We feel that the generalization of specific observations in an experimental framework is an important mechanism for the advancement of theory. Similarly we feel that implementing abstactions frequently leads us to consider important detail that might otherwise be passed over by a strictly theoretical approach. Consequently we have expended great effort on this project because we feel the resulting system is a essential tool for studying the problems inherent in big, heterarchical systems. Understanding such systems seems central at this moment to the continued rapid development of A. I., and therefore our system, being the only laboratory system of its kind, is of considerable importance. Using it allows us to experiment and confirm to a degree impossible without it.

What we use and why

A basic tenet of our approach is that a minimal system, however flimsy, is essential to the effort. With a properly designed, expansion oriented system, we have a tool that greatly promotes the achievement of our general goals, both from the point of view of experimental flexibility and from that of good moral. Without such a minimal system our vision laboratory would be incomplete, like a chemistry lab without test tubes. Our thinking would not be as sound and our results would not be so convincing.

To get our minimal system together, we have tried to use existing methods insofar as possible. Our idea has been to rewrite the programs of Binford, Mahabala, Guzmán, and Winston in such a way that the results work well together and can support future developement. But while tangible progress has been our goal, this is not to say that theoretical advances have been suppressed. Better ideas are to be expected from implementation efforts of this kind. In particular some heterarchical structure has emerged already, some flaws in previous thinking have been exposed, a number of new programs have been written, and some new ideas of the global sort are developing. In particular, Freuder's paper suggests a plausible approach to developing theoretical foundations for much of the region amalgamation work.

The modules

1. The Feature Finder. <Binford>.

   This module performs horizontal and vertical scans. Three scan lines are in core at any given time. The effort is to initiate new lines and to continue already existing lines through the narrow band defined by the three scan lines.

2. The Drawing Drawer <Horn>.

Horn's program does the difficult job of transforming Binford's kruft into a decent line drawing. Lists of related feature points are generated, lines are formed from them, and vertices are concocted at intersections.

3. The Bookkeeper <Winston>

This module creates a PLANNER data base and does vertex identification. It resembles Mahabala's program, but it goes beyond that program by way of complaining about certain highly convex regions.

4. The Proposer <Freuder.>

Freuder distilled what we know of the Horn-Binford program's skill into a line proposer. It appears to be very conservative and rarely proposes a line that is not there.

5. The Verifiers <Lerman,Binford, Herskovits, and Griffith>

Binford and Herskovits left behind a verifier in MIDAS and LISP. Lerman has nicely improved it, debugged it, and purged the LISP part. We are experimenting with it and will have contests between it and Griffith's.

6. The Body Finder <Freuder>.

Freuder has thoroughly reworked Guzman's ideas, with influence from some others, into a far more satisfying body finder. It outperforms Guzman's SEE and is superior esthetically.

7. The Structure Describer <Winston>

This program embodies the SUPPORTED-BY ideas of Winston's thesis. Many of the ideas only suggested there are now implemented. Its results are used by the position locater.

8. The Position Locater <Winston>

This small package uses support information together with calibration results from Horn's calibration routine to fix the position of corners. (This program exhibits some LISP code by Horn, thought to be a rarity.)

9. The Unobscured-pickupable-Brick-Finder <Freuder>

   This is a conservative specialist that works together with the Position Locater to establish the proper arm coordinates for grasping.

10. The Skeleton Extractor <Winston>

   SKELETON is a program that examines bricks thoroughly to establish their dimensions in spite of visual obstructions. It is more general but less sure-footed than Freuder's program above, but it is similarly charged with supplying the Position Locater with good points. (A small routine called FLESH by Freuder interfaces the rather big SKELETON program with the Position Locater.

11. The Copy Planner <Winston>

   This elementary program establishes the sequence of grasp and ungrasp operations adequate for configurations lacking annoying unstable substructures. Peculiarly it first imagines it is taking the structure apart and then reverses the resulting plan.

12. The Free Space Finder <Freuder>

   This supplies the Copy Planner with a location in the storage area called HELL, at which a spare part may temporarily reside before use in the WORLD section of the UNIVERSE, which of course is the black felt-covered table.

13. The SLAVE <Silver and Horn>

   This MIDAS code gets the AMF arm from point A to point B. It tries to smooth out the motion insofar as possible and does the opening, closing, rotating, interrupting, and other chores needed to avoid chaos and danger. Silver's code handles the motion while Horn's does the interfacing and handles the details of organizing the various motions.

14. The Calibrator <Horn>

   This goes through the hair of establishing where the eye is and supplying other routines with a transformation matrix and details of the table's position. A pie shaped black and white object is tracked in the course of this operation.

What we want to do

We want to address important issues through specific, incremental expansion and improvement. Being modular, our system encourages this. Many of the possibilities to be mentioned bear on more than one global question and the categorization below is recognized as loose. We especially welcome criticism and suggestions on the subject of which of the following deserve priority and more thought.

Heterarchy

Certainly the study of how a big, knowledgeable system can work is a major goal. To this end we have in mind channels that carry advice, complaint, and conjecture, as well as those traditionally thought of as data pipes. So far our system has only a few examples of heterarchy, as our prime purpose has been to create a full system, albeit thin.

Vision flashes 7 and 8 outline our past forrays into this area in the /region criticize/line proposer/line verifier/ complex and the /object recognizer/K joint resolver/region conglomerator/ chain. All our new additions will contribute to this study, but it is convenient here to mention the following:

> Focus

Fixing position by Horn's focus routine nicely complements the known-supporting-plane method, providing a great heterarchy situation to talk about and experiment with.

> Stereo

The work of Lerman in his master's thesis raises interesting questions and shows that the matching problem can often be avoided, given objects with texture.

> Identification

Implementation of more of the known ideas for determining object identities will provide other heterarchy hooks.

Generalists and Specialists

It would seem that a system that can cope with the world must have both special and general knowledge. For example, in region conglomeration and in dimension calculation we already have good results blending together programs that know about unobscured bricks with those that have more general ability. To further understand the interaction possible using both sorts of knowledge, the following may be useful:

> Wedge Specialists

Many of our analysis modules work on more or less arbitrary solids that are perpendicular projections of some two-dimensional shape. Others, like the skeleton program, are specialists and are so far limited to bricks. Given a skeleton program for wedges, we will have two specialists doing the same kind of task but in different circumstances.

> Degeneracies

We want to understand bodies seen end on. In all but trivial cases this seems to require deduction and heterarchy as well as special purpose knowledge.

> Color

An unexplored area for us, but one with potential in view of the questions color vision psycology has raised.

Theory of Scenes

Working with a real system inevitably raises theoretical questions and suggests approaches to them. Eugene Freuder's ideas as outlined in Vision Flashes 4 and 5, were favorably influenced by the blend of theoretical and practical factors in our working environment. We are anxious to stimulate more of this and suggest a number of possiblities:

> Shadows

Papert has argued for the development of a shadow theory and Waltz has told me of some interesting ideas on this subject. For example, he notes that a simple algorithm can often establish the position of the light source.

> Texture

> Objects with Curved Surfaces

> Vertex Finding

Lerman has an exciting set of ideas about circular scan. Rather than horizontal and vertical bands which we now use, we would use rings of 3 or 4 circular scans. A program would attempt to shrink these rings onto vertexes. If lines entering a particular ring do not converge suitably on a single point, more rings would be thrown off, one for each apparent convergence point.

Environmental Interaction

Simon claims in The Sciences of the Artificial that much of what appears to be high intelligence is a happy amalgum of simple organism and complex environment. While the point he makes is certainly arguable, surely the interaction of intelligence with the environment deserves some experimental investigation and we are in a unique position to do it. Our intention is to begin with these:

> Performance Monitoring

The robot should look at what it does and resolve any anomalies between what it sees and what it thinks it has done. Work on this is in progress.

> Disassembly

The robot should know when it is beaten and cannot analyze a structure completely. Then it should remove the objects causing the confusion and go on.

> Experiment

If the robot is not sure whether something is one object or two abutting, it should reach out and test.

Representation

We must worry about data base design for heterarchical systems working in a changing world. There are many questions of data consistancy and data exchange between modules to be considered. We want appropriate information available and we want an uncluttered memory. We want descriptions that themselves are part of the solutions to problems like that of locating object storage space.

Miscellaneous

A number of other issues have and will emerge in addition to those discussed above:

> Findspace

This is the problem of finding storage space for an object. There are reasonable methods now for testing particular places. Susman has some nice code for this waiting. But we do not yet know how to find places that are good to test.

> Construction

A construction program can know about counterweights, glue, scaffolds, tools and subassemblies. Fahlman is already at work on such a program.

> PLANNER

As an aside, perhaps we should also note that we are the heaviest users of micro-PLANNER, and as such we try to contribute to the language's further development by way of applause, suggestion, and complaint.

How we want to do it

Organization

It is time to review the organization of the vision project and perhaps redefine its purpose and scope.

The responsibility of the group's various members and the inherent limits on the group's size and budget could be clarified.

The members of the group would prefer to have a hand in decision making with respect to new personnel and equipment.

People

Many A. I. people are interested in robotics, and it is desirable to have at least two available styles of involvement. Graduate students and random, occasionally interested people should contribute when moved, but it is neither fair nor desirable to rely on them for essential work as other demands on their time take preference. In another group, the sort of core robotics group, one should have individuals who can thrive in results mode, the atmosphere being like that maintained in the vision group in the recent past.

We speculate that graduate students will produce new ideas on the more global, abstract, unexplored questions with which our laboratory has neglIgeable experience. Paid shorter-term undergraduates, should concentrate on the seemingly straightforward implementations, which generally prove to unleash plenty of problems to be of educational and theoretical benefit. Salaried graduate students and staff are in an ideal position to both abstract theory from results and realize theory as results. Consequently they should mix their efforts so that they not only keep close to the system and improve it, but also bring their experience to bear on the deeper theoretical issues that emerge and become difficult long term problems.

The core group needs Freuder, Lerman, Griffith, Winston, and a Horn substitute. The intention is to wait patiently and be very selective in attempting to fill vacant slots. Any person should at least potentially be available for beyond one year's time. We would like to have two or more summer student-type employees if suitable candidates appear by way of work in 6.258 or 6.544 term papers. The extra supervision time available in the summer permits good use of such people.