What is an Image?

Harold Cohen
University of California at San Diego
B-027, UCSD, La Jolla, CA 92093

Image-making, and more particularly art-making, are considered as rule-based activities in which certain fundamental rule-sets are bound to low-level cognitive processes. AARON, a computer-program, models some aspects of image-making behavior through the action of these rules, and generates, in consequence, an extremely large set of highly evocative "freehand" drawings. The program is described, and examples of its output given. The theoretical basis for the formulation of the program is discussed in terms of cultural considerations, particularly with respect to our relationship to the images of remote cultures. An art-museum environment implementation involving a special-purpose drawing device is discussed. Some speculation is offered concerning the function of randomizing in creative behavior, and an account given of the use of randomness in the program. The conclusions offered bear upon the nature of meaning as a function of an image-mediated transaction rather than as a function of intentionality. They propose also that the structure of all drawn images, derives from the nature of visual, cognition.

1. INTRODUCTION

AARON is a computer program designed to model some aspects of human art-making behavior, and to produce as a result "freehand" drawings of a highly evocative kind (figs 1,2). This paper describes the program, and offers in its conclusions a number of propositions concerning the nature of evocation and the nature of the transaction – the making and reading of images - in which evocation occurs. Perhaps unexpectedly – for the program has no access to visual data – some of these conclusions

bear upon the nature of visual representation. This may suggest a view of image-making as a broadly referential activity in which various differentiable modes, including what we call visual representation (note 1), share a significant body of common characteristics.

in some respects the methodology used in this work relates to the modeling of "expert systems" (note 2), and it does in fact rely heavily upon my own "expert" knowledge of image-making. But in its motivations it cones closer to research in the computer simulation of cognition. This is one area, I believe, in which the investigator has no choice but to model the human prototype. Art is valuable to human beings by virtue of being made by other human beings, and the question of finding more efficient modes than those which characterize human performance simply does not arise.

My expertise in the area of image-making rests upon many years of professional activity as an artist — a painter, to be precise (note 3) — and it will be clear that my activities as an artist have continued through my last ten years of work in computer-modeling. The motivation for this work has been the desire to understand more about the nature of art-making processes than the making of art itself allows, for under normal circumstances the artist provides a near-perfect example of an obviously-present, but virtually inaccessible body of knowledge. The work has been informal, and qua psychology lacks methodological rigor. It is to be hoped, however, that the body of highly specialized knowledge brought to bear on an elusive problem will be some compensation.

AARON is a knowledge-based program, in which knowledge of image-making is represented in rule form. As I have indicated I have been my own source of specialized knowledge, and I have served also as my own knowledge-engineer. before embarking on a detailed account of the program's workings, I will describe in general terms what sort of program it is, and what it purports to do.

First, what it is 'not. It is not an "artists' tool". I mean that it is not interactive, it is not designed to implement key decisions made by

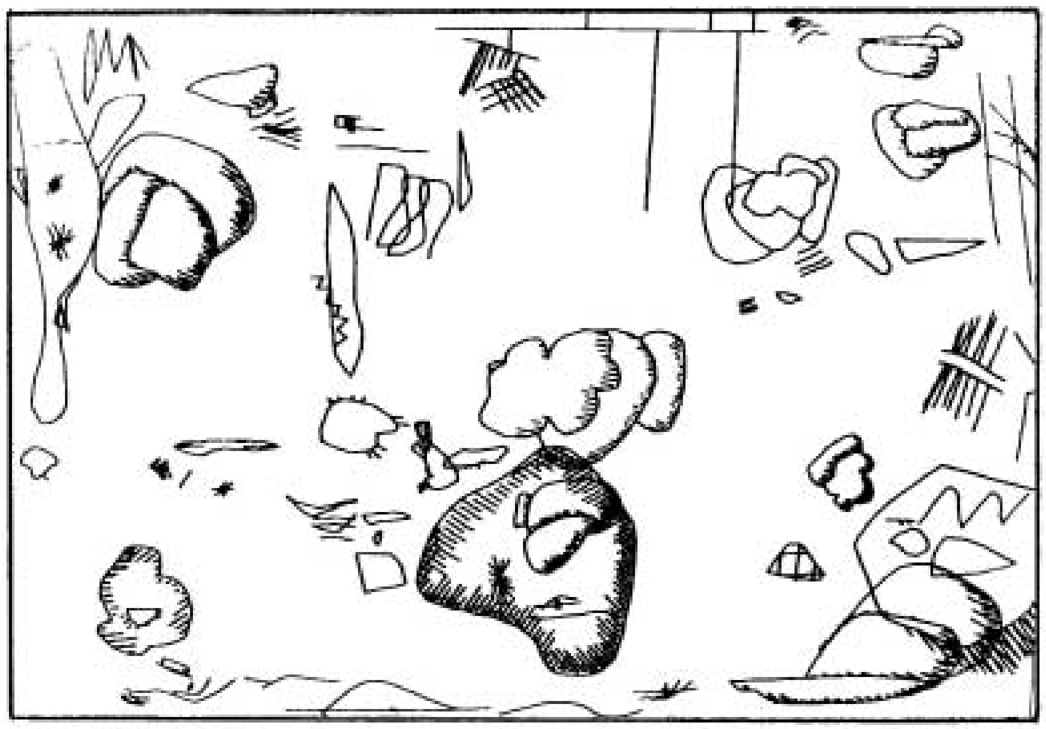


figure 1.

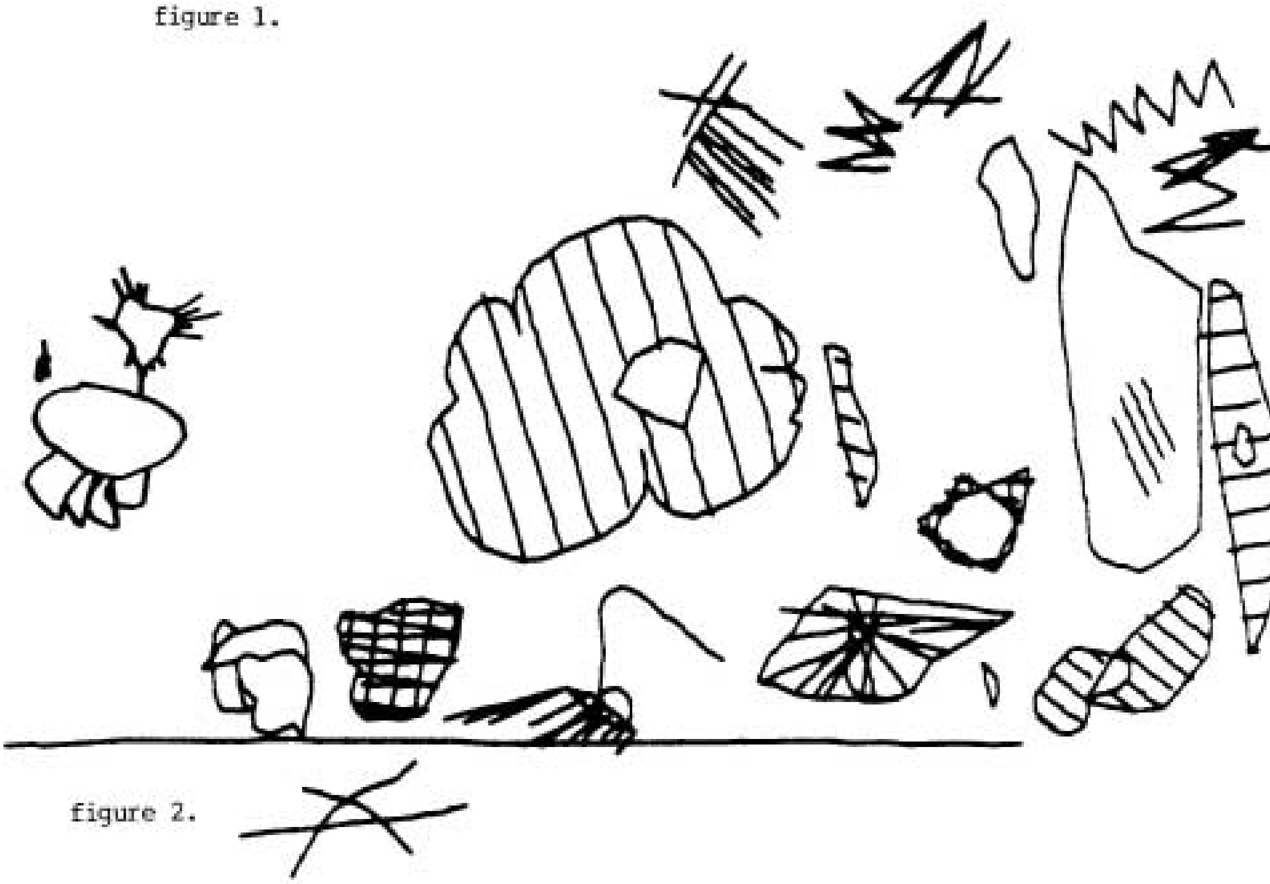


figure 2.

the user, and it does not do transformations upon input data. in short, it is not an instrument, in the sense that most computer applications in the arts, and in music particularly, have identified the machine in essentially instrument-like terms.

AARON is not a transformation device. There is no input, no data, upon which transformations could be done: in fact it has no data at all which it does not generate for itself in making its drawings. There is no lexicon of shapes, or parts of shapes, to be put together, assembly line fashion, into a complete drawing.

It is a complete and functionally independent entity, capable of generating autonomously an endless succession of different drawings (note 4). The program starts each drawing with a clean sheet of paper — no data — and generates everything it needs as it goes along, building up as it proceeds an internal representation of what it is doing, which is then used in determining subsequent developments. It is event driven, but in the special sense that the program itself generates the events which drive it.

It is not a learning program, has no archival memory, is quite simple and not particularly clever. It is able to knock off a pretty good drawing — thousands, in fact — but has no critical judgment that would enable it to declare that one of its drawings was "better" than another. That has never been part of the aim. Whether or not it might be possible to demonstrate that the artist moves towards higher goals, and however he might do so through his work, art-making in general lacks clear internal goal-seeking structures. There is no rational way of determining whether a "move" is good or bad the way one might judge a move in a game of chess, and thus no immediately apparent way to exercise critical judgment in a simulation.

This lack of internal goal-orientation carries with it a number of difficulties for anyone attempting to model art-making processes: for one thing, evaluation of the model must necessarily be informal. In the case of AARON, however, there has been extensive testing. Before describing the testing procedure it will be necessary to say with more care distinguishing here between the program's goals and my own — what AARON is supposed to do.

Task Definition.

It is not the intent of the AARON model to turn out drawings which are, in some ill-defined and loosely-understood sense, aesthetically pleasing, though it does in practice turn out pleasing drawings. It is to permit the examination of a particular property of freehand drawing which I will call, in a deliberately general fashion, standing-for-ness.

The Photographic "Norm"

One of the aims of this paper is to give clearer definition to what may seem intuitively obvious about standing-for-ness, but even at the outset the "intuitively obvious" will need to be treated with some caution, in particular, we should recognize that unguarded assumptions about the nature of "visual" imagery are almost certain to be colored by the XXth century's deep preoccupation with photography as the "normal" image-making mode. The view that a drawn image is either:

1. representational (concerned with the appearance of things) , or

2. an abstraction (i.e. fundamentally appearance-oriented, but transformed in the interest of other aims) or,

3. abstract (i.e. it doesn't stand for anything at all),

betrays just this pro-photographic filtering, and is a long way from the historical truth. There is a great wealth of imagistic material which fits none of these paradigms, and it is by no means clear even that a photograph carries its load of standing-for-ness by virtue of recording the varying light intensities of a particular view at a particular moment in time.

It is for this reason that image-making will be discussed here as the set of modes which contains visual representation as one of its members. It is also why I used the word "evocative" in the first paragraph rather than "meaningful". My domain of enquiry here is not the way in which particular meanings are transmitted through images and how they are changed in the process, but more generally the nature of image-mediated transactions. What would be the minimum condition under which a set of marks may function as an image? This question characterizes economically the scope of the enquiry, and it also says a good deal about how the word "image" is to be used in this paper, though a more complete definition must wait until the end.

Art-making and Image-making.

The reader may detect some reluctance to say firmly that this research deals with art-making rather than with image-making, or vice-versa. The two are presented as continuous. Art-making is almost always a highly sophisticated activity involving the interlocking of complex patterns of belief and experience, while in the most general sense of the term image-making appears to be as "natural" as talking. All the same, art-making is a case of image-making, and part of what AARON suggests is that art-making rests upon cognitive processes which are absolutely normal and perfectly common.

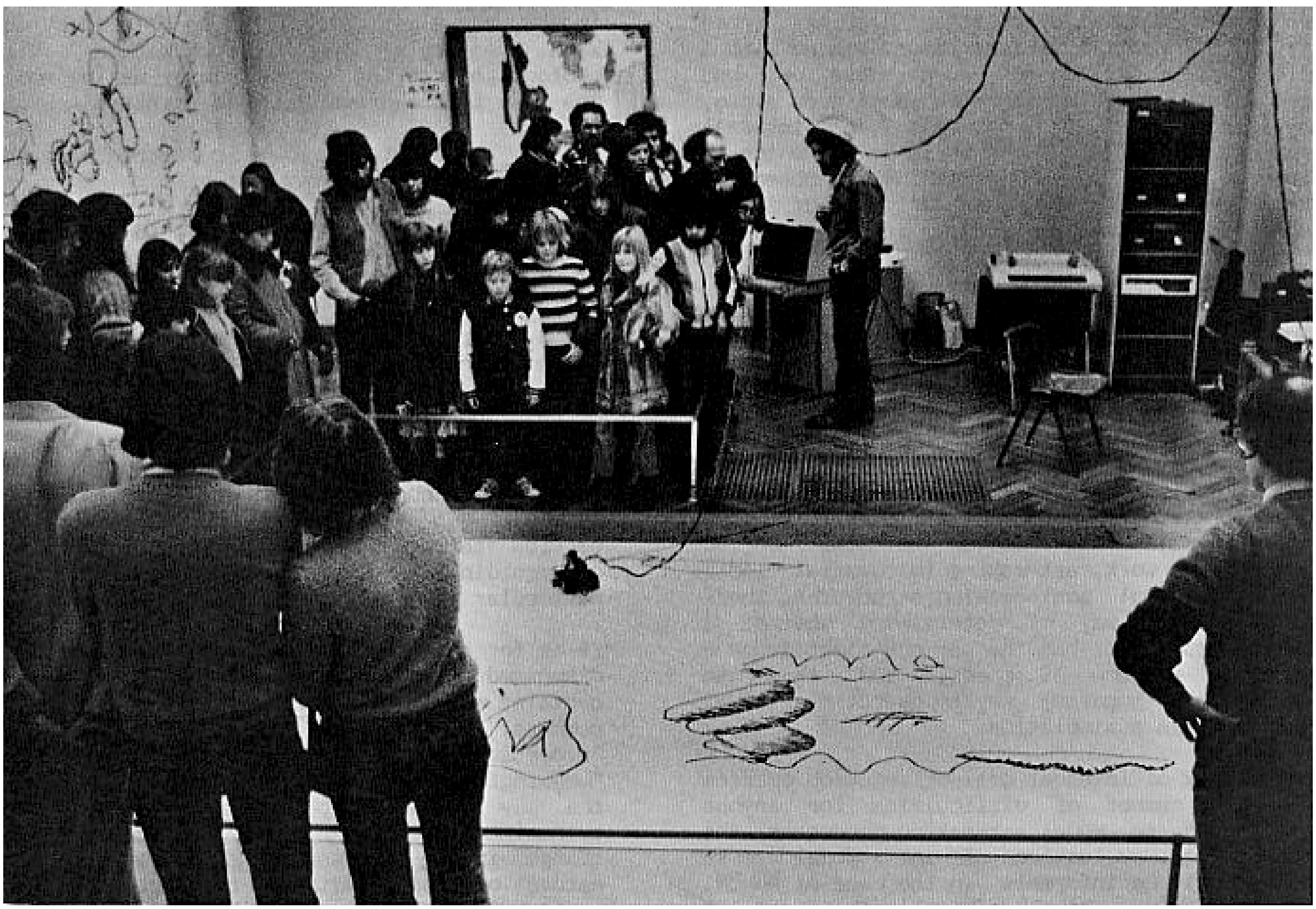


figure 3.

Evaluation.

A simulation program models only a small piece of the action, and it requires a context in which to determine whether it functions as one expects that piece to function. AARON is not an artist. It simply takes over some of the functions which normally require an artist to perform them, and thus it requires the whole art-making process to be carried forward as a testing context. The program's output has to be acceptable to a sophisticated audience on the same terms as any other art, implying thereby that it must be seen as original and of high quality, not merely as a pastiche of existing work.

A valid testing procedure must therefore contain a sophisticated art-viewing audience, and the informal in situ evaluation of the simulation has been carried out in museum environments: the DOCUMENTA 6 international exhibition in Kassel, Germany, and the prestigious Stedelijk Museum in Amsterdam, the two exhibits together running for almost five months and with a total audience of almost half a million museum-goers, in both of these shows drawings were produced continuously on a Tektronix 4014 display terminal and also with an unconventional hard-copy device (to be described later) . A PDP 11/34 ran the program in full view of the gallery visitors (fig 3).

In addition and at other times the program's output has been exhibited in a more orthodox mode in museums and galleries in the US and in Europe.

These exhibits were not set up as scientific experiments. Nor could they have been without distorting the expectations of the audience, and thus the significance of any results. No formal records were kept of the hundreds of conversations which took place between the artist and members of the audience. This report is therefore essentially narrative, but offered with some confidence.

Audience Response.

A virtually universal first assumption of the audiences was that the drawings they were watching being made by the machine had actually been made in advance by the "real" artist, and somehow "fed" to the machine. After it had been explained that this was not the case viewers would talk about the machine as if it were a human artist. There appeared to be a general consensus that the machine exhibited a good-natured and even witty artistic personality, and that its drawings were quite droll (fig 4). Some of the viewers, who knew my work from my pro-computing, European, days claimed that they could "recognize my hand" in the new drawings. This last is particularly interesting, since, while I certainly made use of my own body of knowledge concerning image-making in writing the program, the appearance of my own work never consciously served as a model for what the program was supposed to do.

More to the point, while a very small number of people insisted that the drawings were nothing but a bunch of random squiggles, the majority clearly saw them in referential terms. Many would stand for long periods watching, and describing to each other what was being drawn; always in terms of objects in the real world. The drawings seem to be viewed mostly as landscapes inhabited by "creatures", which would be "recognized" as animals, fish, birds and bugs. Occasionally a viewer would "recognize" a landscape, and once the machine's home was identified as San Francisco, since it had just drawn Twin Peaks.

It might be correctly anticipated that on those other occasions when drawings have simply been framed and exhibited without any reference to their origins, the question of their origins has never arisen, and they have met with a typical cross-section of museum-goers responses.

Even without formal evaluation, it might reasonably be claimed that the program provides a convincing simulation of human performance.

The next part of this paper is divided into five sections. In the first, a general description of the production system as a whole is given. The following three sections deal with particular parts of the production system: the MOVEMENT CONTROL part, the PLANNING part, and the part which handles the internal representation of the drawings as they proceed. The second of these, on PLANNING, also gives an account of the theoretical basis for the program. The fifth section has something to say about the function of randomness in the program, and also discusses to what extent it might be thought to parallel the use of randomness in human art-making behavior. The third and final part draws conclusions.

****

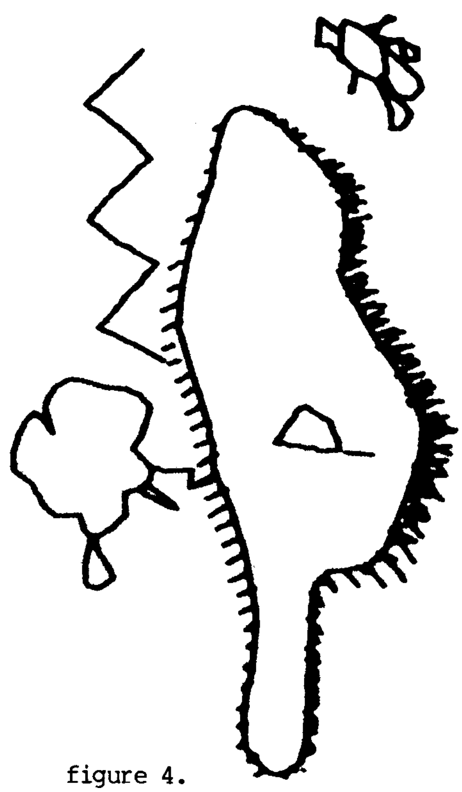


figure 4.

2. THE PROGRAM "AARON"

2.1 THE PRODUCTION SYSTEM.

The main program (note 5) has about three hundred productions. Many of these are to be regarded as micro-productions in the sense that each of them handles only a small part — an "action-atom" — of a larger conceptual unit of action. For example, the drawing of a single line, conceptually a single act, actually involves twenty or thirty productions on at least three levels of the system. This fine-grain control over the drawing process subscribes both to its generality — most of these action-atoms are invoked by many different situations — and to its flexibility, since it allows a process to be interrupted at any point for further consideration by higher-level processes.

Levels of Organization.

The organization of the system is hierarchical, in the sense that the higher levels are responsible for decisions which constrain the domain of action for the lower levels (fig 5). Each level of the system is responsible only for its own domain of decision-making, and there is no conceptual homunculus sitting on the top holding a blueprint and directing the whole operation. No single part knows what the drawing should turn out to be like. There is some practical advantage to this level-wise splitting up of the system, but the program was designed this way primarily for reasons of conceptual clarity, and from a desire to have the program structure itself — as well as the material contained within it — reflect my understanding of what the human image-making process might be like. I believe that the constant shifting of attention to different levels of detail and conceptualization provides this human process with some of its important characteristics. Thus the left part of each production searches for combinations of up to five or six conditions, and each right part may perform an arbitrary number of actions or action-atoms, one of which may involve a jump to another level of the system.

"ARTWORK"

The topmost level of the system, the ARTWORK level, is responsible for decisions relating to the organization of the drawing as a whole. It decides how to start, makes some preliminary decisions which may later determine when and how it is to finish, and eventually makes that determination. The program currently has no archival memory, and begins each drawing as if it has never done one before. (One can easily imagine the addition of a higher level designed to model the changes which the human artist deliberately makes in his work from one piece to the next; this level would presumably be called EXHIBITION.)

ARTWORK also handles some of the more important aspects of spatial distribution. It is my belief that the power of an image to convince us that it is a representation of some feature of the visual world rests in large part upon the image's fine-grain structure: the degree to which it seems to reflect patterns in the changes of information density across the field of vision which the cognitive processes themselves impose upon visual experience.

Put crudely, this means, for example, that a decision on the part of the reader of an image that one set of marks is a detail of another set of marks rather than standing autonomously, is largely a function of such issues as relative scale and proximity. This function is quite apart from the more obviously affective issue of shape (and hence "semantic") relationship, it is the overall control of the varying density of information in the drawing, rather than the control of inter-figural relationships, which is handled by ARTWORK.

"MAPPING" and "PLANNING"

All problems involving the finding and allocation of space for the making of individual elements in the drawing is handled by MAPPING, though its functions are not always hierarchically higher than those of PLANNING, which is responsible for the development of these individual figures. Sometimes PLANNING may decide on a figure and ask MAPPING to provide space, while at other times MAPPING may announce the existence of a space and then PLANNING will decide what to do on the basis of its availability. Sometimes, indeed, MAPPING may override a PLANNING decision by announcing that an appropriate space is not available. A good example of this occurs when PLANNING decides to do something inside an existing closed figure and MAPPING rules that there isn't enough room, or that what there is the wrong shape.

MAPPING will be referred to again in relation to the data-structures which constitute the program's internal representation of what it is doing, and PLANNING also as one of the centrally important parts of the program.

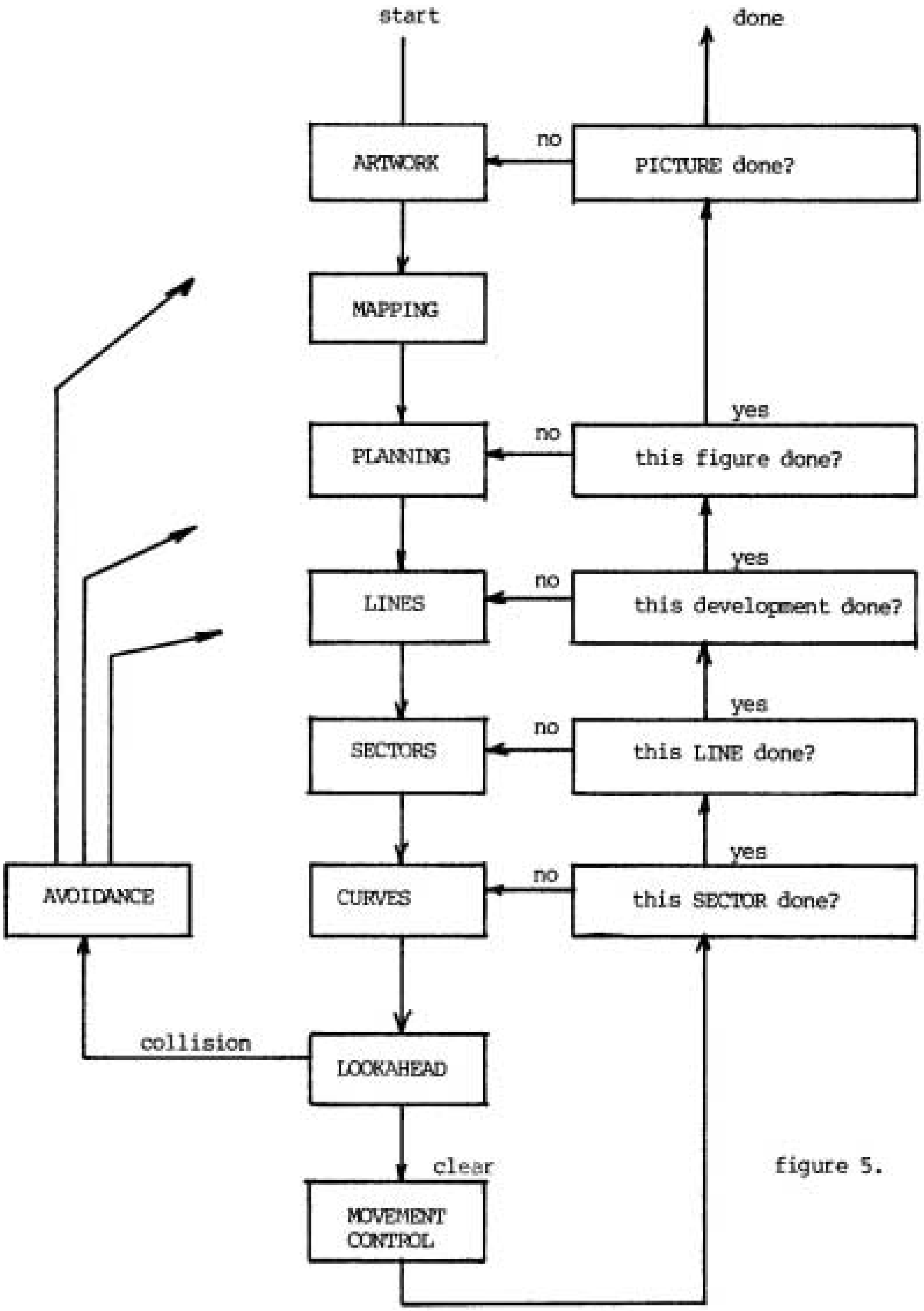

start
done
ARTWORK
no
PICTURE done?
MAPPING
yes
no
PLANNING
this figure done?
yes
no
LINES
this development done?
yes
no
SECTORS
this LINE done?
yes
no
AVOIDANCE
CURVES
this SECTOR done?
collision
LOOKAHEAD
clear
MOVEMENT CONTROL

figure 5.

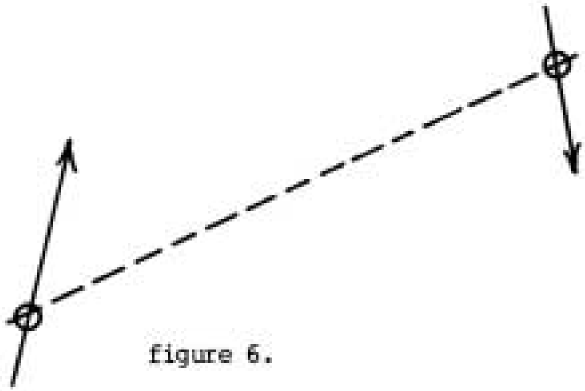


figure 6.

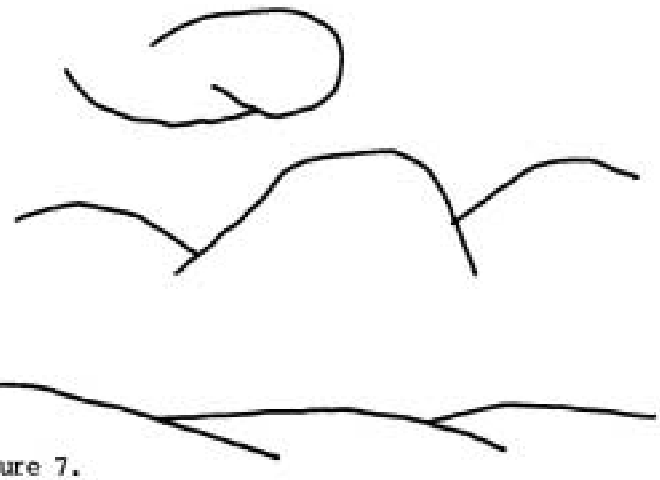


figure 7.

LINES AND SECTORS

Below this level the hierarchical structure of the system is fairly straightforward. Each figure is the result of (potentially) several developments, each provided by a return of control to PLANNING. Each of these developments may consist of several lines, and for each of the successive lines of each development of any figure LINES must generate a starting point and an ending point, each having a direction associated with it (fig 6). It also generates a number of parameters on the basis of specifications drawn up in PLANNING which

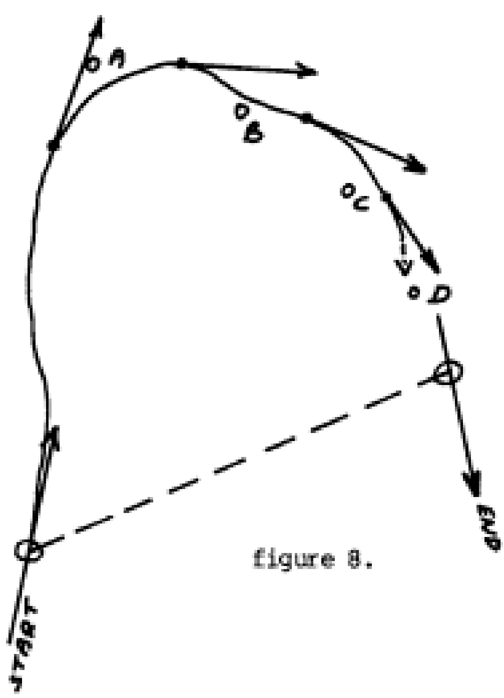


figure 8.

determine how the line is to be drawn: whether reasonably straight, wiggly, or strongly curved, and, if various overlapping modes are called for (fig 7), how they are to be handled.

As I have indicated, lines are not drawn as the result of a single production. When LINES passes control to SECTORS the program does not know exactly where the line will go, since the constraint that it must start and end facing specified directions does not specify a path: there are an indeterminate number of paths which would satisfy the constraint. The program does not choose one, it generates one. SECTORS produces a series of "imagined" partial destinations — signposts, as it were (fig 8) — each designed to bring the line closer to its final end-state. On setting up each of these signposts it passes control to CURVES, whose function is to generate a series of movements of the pen which will make it veer towards, rather than actually to reach, the current signpost. Control is passed back to SECTORS when the pen has gone far enough towards the current signpost that it is time to look further ahead, and it is passed back to LINES when the current line has been completed and a new one is demanded by the development still in progress.

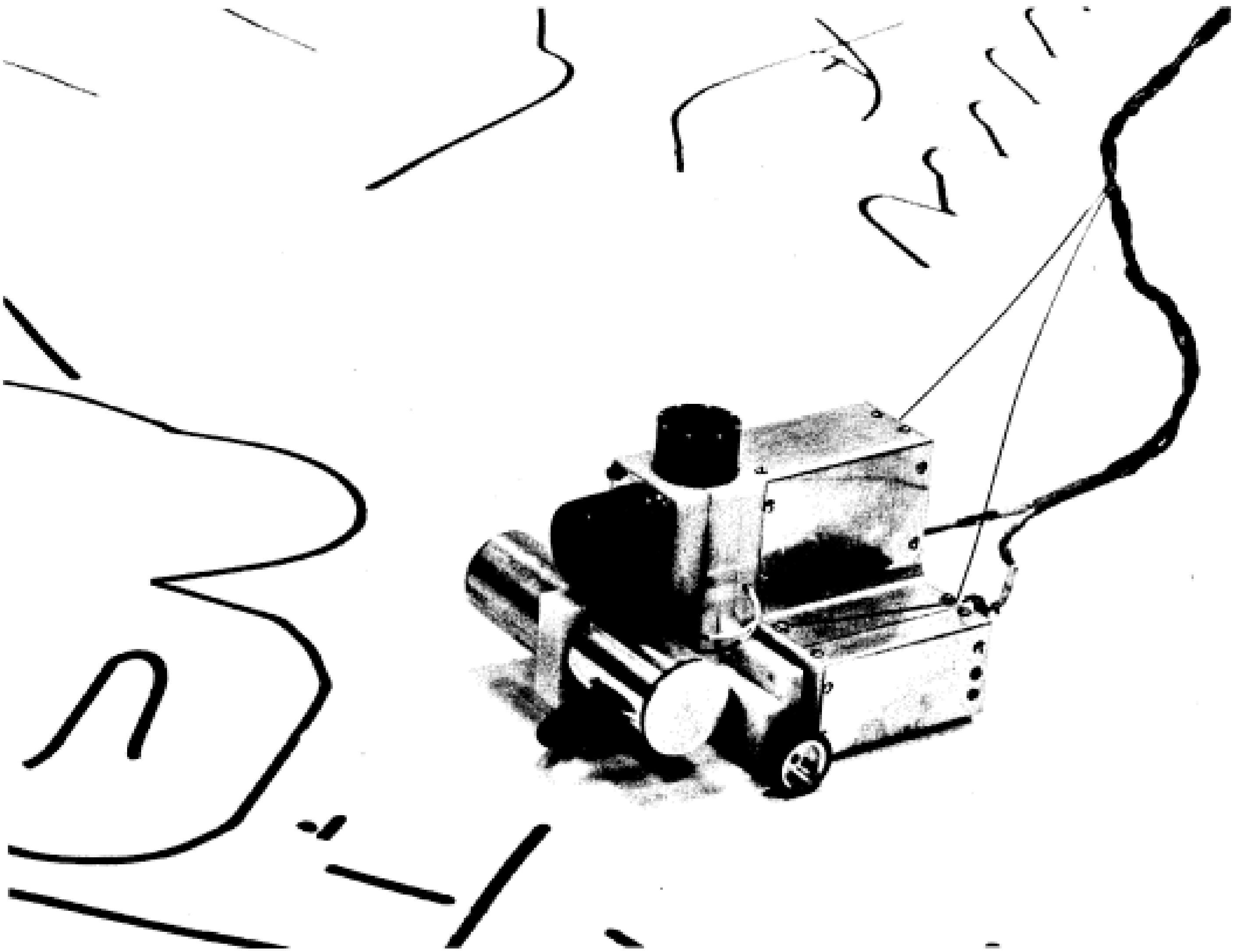
2.2 MOVEMENT CONTROL

We are now down to the lowest level of the program, and to the way in which the curves which make up the drawing are actually generated. This part is not discontinuous from the rest, of course. The flexibility of the program rests in large part upon the fact that the hierarchy of control extends downwards to the finest-grained decisions: no part of the control structure is considered to be so automatic that it should fall below the interface line. Thus, the story of how the pen gets moved around follows on from the description of how the intermediate signposts are set up.

Abstract Displays and Real Devices

In the earlier versions of the program all the development work was done exclusively on a graphic display, and the "pen" was handled as an abstract, dimensionless entity without real-world constraints upon its movements. Conceptually, however, I always thought of the problem of moving the pen from point A facing direction alpha, to point B facing direction beta, as being rather like the task of driving a car off a main road into a narrow driveway set at some known arbitrary angle to it. This is clearly not a dead-reckoning task for the human driver, but one which involves continuous feedback and a successive-approximation strategy.

It seemed quite reasonable, therefore, to be faced at some point with the problem of constructing an actual vehicle which would carry a real pen and make real drawings on real sheets of paper. That situation arose in the Fall of '76 when I was preparing to do the museum exhibitions which I mentioned earlier, and decided that if I wanted to make the drawing process visible to a large number of people simultaneously, I would need to use something a good deal bigger than the usual graphic display with its 20-inch screen.



The Turtle.

The answer turned out to be a small two-wheeled turtle (fig 9), each of its wheels independently driven by a stepping motor, so that the turtle could be steered by stepping the two motors at appropriate rates. It is thus capable of drawing arcs of circles whose radius depends upon the ratio of the two stepping rates.

Since the two wheels can be driven at the same speed in opposite directions, the turtle can be spun around on the spot and headed off in a straight line, so that this kind of device is capable of simulating a conventional x-y plotter. But it seemed entirely unreasonable to have built a device which could be driven like a car and then use it to simulate a plotter. In consequence the pen-driving procedures already in the program were re-written to generate the stepping rates for the motors directly — to stay as close as possible to the human model's performance — rather than calculating these rates as a function of decisions already made.

The advantage here was a conceptual one, with some practical bonus in the fact that the turtle does not spend a large part of its time spinning instead of drawing. It also turned out unexpectedly that the generating algorithm simplified enormously, and the quality of the freehand simulation improved noticeably.

Feedback.

The program does not now seek to any place — in Cartesian terms — but concerns itself exclusively with steering: thus the turtle's Cartesian position at the end of executing a single command is not known in advance. Nor is this calculation necessary when the turtle is operating in the real world. It was not designed as a precision drawing device, and since it cannot perform by dead-reckoning for long without accumulating errors, the principle of feedback operation was extended down into this real-world part of the program, the device makes use of a sonar navigation system (fig 10) by means of which the program keeps track of where it actually is. instead of telling it to "go to x,y" as one would tell a conventional plotter, the program tells it "do the following and then say where you are".

A more detailed account of the turtle system, and it's effect upon the simulation of freehand drawing dynamics, is given in Appendix 1.

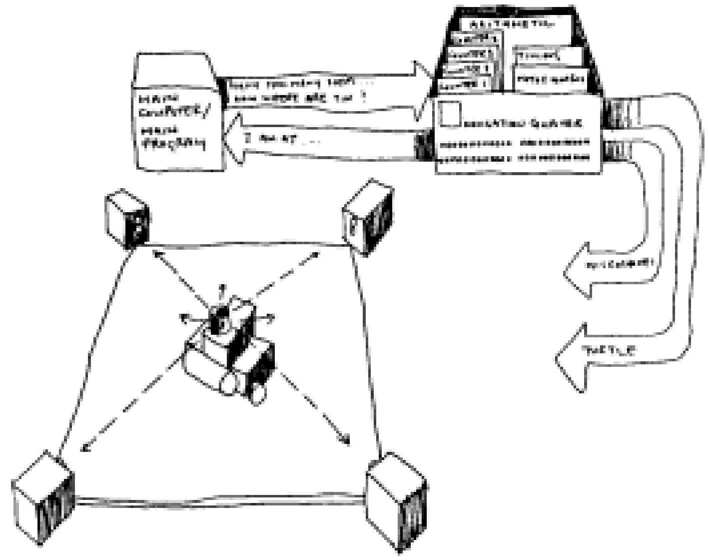


figure 10.

2.3 "PLANNING"

No single level of the program can be described adequately without reference to the other levels with which it interacts: it has already been mentioned, for example, that MAPPING may either precede PLANNING in determining what is to be done next, or it may serve PLANNING by finding a space specified there. Additionally, any development determined in PLANNING may be modified subsequently either as a result of an imminent collision with another figure or because provision exists in the program for "stacking" the current development in order to do something not originally envisaged (fig lla,b). All the same, the drawing is conceived predominantly as an agglomeration of figures, and to that extent PLANNING, which is responsible for the development of individual figures, is of central importance.

Behavioral Protocols in Image-Making.

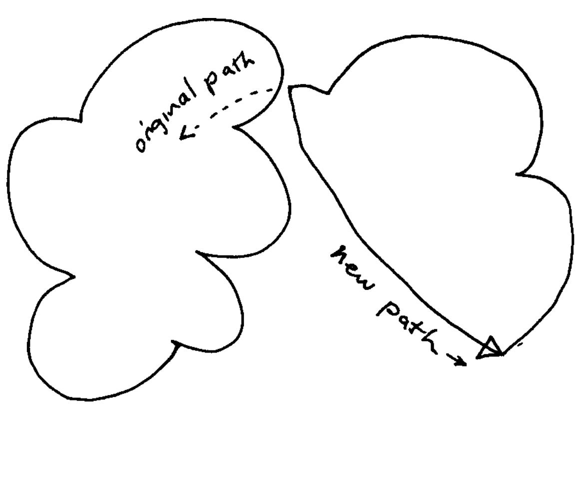
Of the entire program, it is also the part least obviously related to the effects which it accomplishes. While the formal results of its actions are clear enough — an action calling for the closure of a shape will cause it to close, for example — it is not at all clear why those actions result in the specifically evocative quality which the viewer experiences.

A rule-by rule account of this effect is not appropriate, because the individual rules do no more than implement conceptual entities — which I will call behavioral protocols — which are the fundamental units from which the program is built. These protocols are never explicitly stated in the program, but their existence is what authorizes the rules. Thus, before describing in detail what is in PLANNING I should give an account of the thinking which proceeded the writing of the program, and try to make clear what I mean by a protocol.

Background.

It is a matter of fact that by far the greatest part of all the imagery to which we attach the name of "art" comes to us from cultures more or less remote from our own. it is also a matter of fact that within our own culture, and in relation to its recent past, our understanding of imagery rest to a great extent upon prior common understandings, prior cultural agreements, as to what is to stand for what — prior, that is to say, to the viewing of any particular image. It is unlikely that a Renaissance depletion of the Crucifixion ("of Christ" being understood here by means of just such an agreement !) would carry any great weight of meaning if we were not already familiar both with the story of Christ and with the established conventions for dealing with the various parts of the story, indeed, we might be quite confused to find a depletion of a beardless, curly-headed youth on the cross unless we happened to possess the non-obtuse knowledge that Christ was depicted that way — attaching a new set of meanings to the old convention for the representation of Dionysus — until well into the 7th century, in general, we are no longer party to the agreements which make this form acceptable and understandable. We must evidently distinguish between what is understandable without abstruse knowledge — we can, indeed, recognize the figure on the cross as a figure — and what is understandable only by virtue of such knowledge.

In the most general sense, all cultural conditions are remote from us, and differ only in the degree of their remoteness. We cannot really comprehend why the Egyptians made sphinxes, what Michelangelo thought the ancient world had in common with Christianity, or how the internal combustion engine was viewed by the Italian Futurists seventy years ago who wanted to tear down the museums in its honor. What abstruse knowledge we can gain by reading Michelangelo's writings, or the Futurist Manifesto, does not place us into the cultural



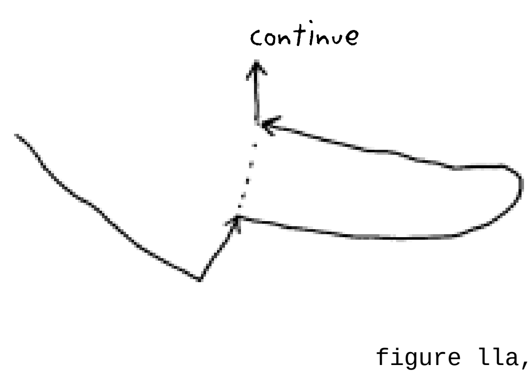


figure lla,b

environment in which the work is embedded. A culture is a continuum, not a static event: its understandings and meanings shift constantly, and their survival may appear without close scrutiny to be largely arbitrary, in the extreme case, we find ourselves surrounded by the work of earlier peoples so utterly remote from us that we cannot pretend to know anything about the people themselves, much less about the meanings and purposes of their surviving images.

The Paradox of insistent Meaningfulness.

There is an implicit paradox in the fact that we persist in regarding as meaningful — not on the basis of careful and scholarly detective work, but on a more directly confrontational basis — images whose original meanings we cannot possibly know, including many that bear no explicitly visual resemblance to the things in the world. Presumably this state of affairs arises in part from a fundamental cultural egocentrism — what, we ask, would we have intended by this image and the act of making it? — which is fundamentally distortive. There has also been a particular confusion in this century through the widespread acceptance of what we might call the telecommunications model of our transactions through imagery, particularly since in applying that model no differentiation has been observed between the culture we live in and the cultures of the remote past. In the view of this model, original meanings have been encoded in the image, and the appearance of the image in the world effects the transmission of the meanings. Allowing for noise in the system — the inevitability of which gives rise to the notion, in art theory, of "interpretation" — the reception and decoding of the image makes the original meanings available.

However useful the model is as a basis for examining real telecommunication-like situations, in which the intended meanings and their transformations can be known and tracked, it provides a general account of our transactions through images which is quite inadequate. The encoding and decoding of messages requires access to the same code-book by both the image-maker and the image-reader, and that code-book is precisely what is not carried across from one culture to another.

I think it is clear also that the paradox of insistent meaningfulness, as we might call it, constitutes the normal condition of image-mediated transactions, not an abnormal condition. It evidently extends below the level at which we can recognize the figure, but not what the figure stands for, since so much of the available imagery is not in any very obvious sense "representational" at all. The paradox is enacted every time we look at a few marks on a scrap of paper and proclaim them to be a face, when we know perfectly well that they are nothing of the sort.

Cognitive Bases for Image Structure.

In short, my tentative hypothesis in starting work on AARON was that all image-making and all image-reading is mediated by cognitive processes of a rather low-level kind, presumably processes by means of which we are able to cope also with the real world. In the absence of cannon cultural agreements these cognitive processes would still unite image-maker and image-viewer in a single transaction. On this level — but not on the more complex culture-bound level of specific iconological intentionality — the viewer's egocentricity might be justified, since he could correctly identify cognitive processes of a familiar kind in the making of the image. But let me detail this position with *some* care. I am not proposing that these processes make it possible for us to understand the intended meanings of some remotely-generated image: I am proposing that the intended meanings of the maker play only a relatively small part in the sense of meaningfulness. That sense of meaningfulness is generated for us by the structure of the image rather than by its content.

I hope I may be excused for dealing in so abbreviated a fashion with issues which are a good deal less than self-evident. The notion of non-enculturated behavior — and that notion lurks behind the last few paragraphs, obviously — is a suspect one, since all human behavior is enculturated to some degree: but my purpose was not to say what part of human behavior is dependent upon enculturating processes and what is not. It was simply to identify some of the determinants to a general image-structure which could be seen to be common to a wide range of enculturating patterns. The implication seemed strong — and still does — that the minimum condition for generating a sense of meaningfulness did not need to include the assumption of an intent to communicate: that the exercise of an appropriate set of these cognitive processes would itself be sufficient to generate a sense of meaningfulness.

Cognitive Skills.

The task then was to define a suitable set. I have no doubt that the options are wide, and that my own choices are not exclusive. I chose at the outset to include:

1. the ability to differentiate between figure and ground,

2. the ability to differentiate between open and closed forms, and

3. the ability to differentiate between insideness and outsideness (note 6) .

AARON has developed a good deal from that starting point, and some of its current abilities clearly reflect highly enculturated patterns of behavior. For example, the program is now able to shade figures in a mode distinctly linked to Renaissance and post-Renaissance representational modes: other cultures have not concerned themselves with the fall of light on the surfaces of objects in the same way. Nevertheless, a large part of the program is involved still in demonstrating its awareness of the more primitive differentiations.

Protocols and Rules.

Against this background, I use the term protocol to mean the procedural instantiation of a formal awareness. This is clearly a definition which rests upon cognitive, rather than perceptual, modes, since it involves the awareness of relational structures. Thus, for example, the program's ability to differentiate between form and ground makes possible an awareness of the spatial relationships between forms, and generates finally a set of avoidance protocols, the function of which is to prohibit the program from ignoring the existence of one figure in drawing another one. The protocols themselves are not explicitly present in the program, and are manifested only through their enactment by the rules which describe what to do in particular circumstances where the overlapping of figures is threatened.

Figure Development

in keeping with the hierarchical structuring which informs the program as a whole, PLANNING considers a figure to be the result of a number of developments, each determined in part by what has gone before. The program enacts a number of repetition protocols, and a single development in the making of a figure can often involve the repetition of a single action (fig 12), rather than the agglomeration of different actions. The first productions to deal with the first development of any figure decide, on the basis of frequency considerations, that this figure will be closed, that it will be open, or that it will be, for the moment, "uncommitted" — that is, a line or a complex of lines will be drawn, but only at a later stage will it be decided tether or not to close. If the primary decision is for closure,

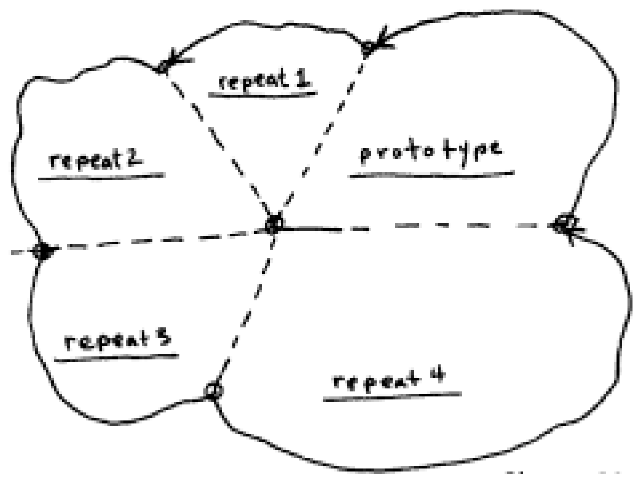


figure 12.

then PLANNING will decide between a number of options, mostly having to do with size and shape — MAPPING permitting — and with configuration, in sane cases it will not actually draw the boundary of a closed form at all, and will leave the definition of the occupied space to await subsequent space-filling moves.

If the decision is for a non-closed form, then again a number of options are open. In both cases the available options are stated largely in terms of repetition protocols, the enactment of which determines the formal characteristics of the resulting configuration. These characteristics are not uniquely defining, however, and a number of different formal sub-groups may result from a single repetition protocol and its rules. For example, one such protocol, involving a single line in this case, requires the line to move a given distance (more-or-less) and then change direction, continuing this cycle a given number of times. All the figures marked in (fig 13) result from this: the details of implementation in the individual cases are responsive to their unique environmental conditions, and in any case may be changed at any point by the overriding avoidance protocol, which guarantees the territorial integrity of existing figures.

Thus the program will know at the beginning of each development what the current intention is, but will not know what shape will result. A closed form generated by a "go, turn, repeat" cycle may in fact turn out to be extremely long and narrow (fig 14), and a number of second developments associated with a closed-form first development will then be unavailable: there will be a limit, for example, upon what can be drawn inside it, though it may develop in other ways, as this one does.

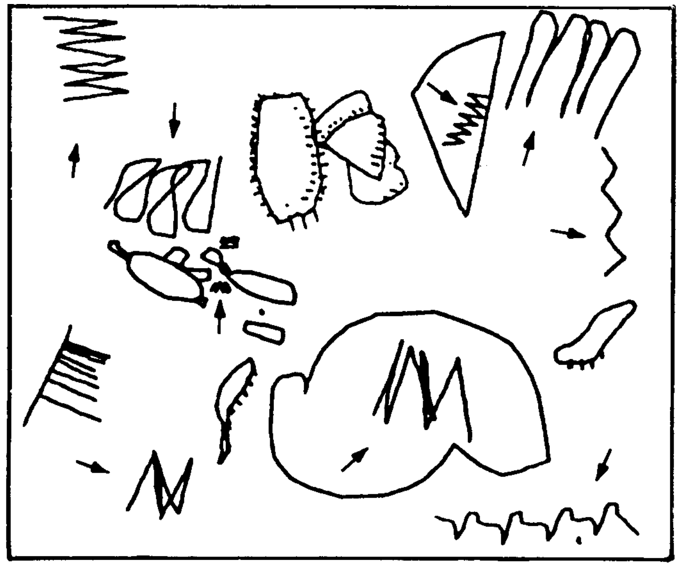


figure 13.

Proliferation.

Even with constraints of this sort there is a significant proliferation in the number of productions associated with the second development of any figure. A typical first development might be initiated by:

If (this is a first development
and the last figure was open
and at least n figures have been done
and at least q of them were open
and at least t units of space are now
available)

Then

This figure will be closed
specifications for repetition
specifications for configuration
to move on from this point:
If (this is a second development
and the first was closed
and its properties were
a. (size)
b. (proportions)
c. (complexity)
d. (proximity to ...)

Then either

1. divide it
... specifications ...
or
2. shade it
... specifications...
or
3. add a closed form to it
... specifications...
4. do a closed form inside
... specifications...
or
5. do an open form inside
... specifications...

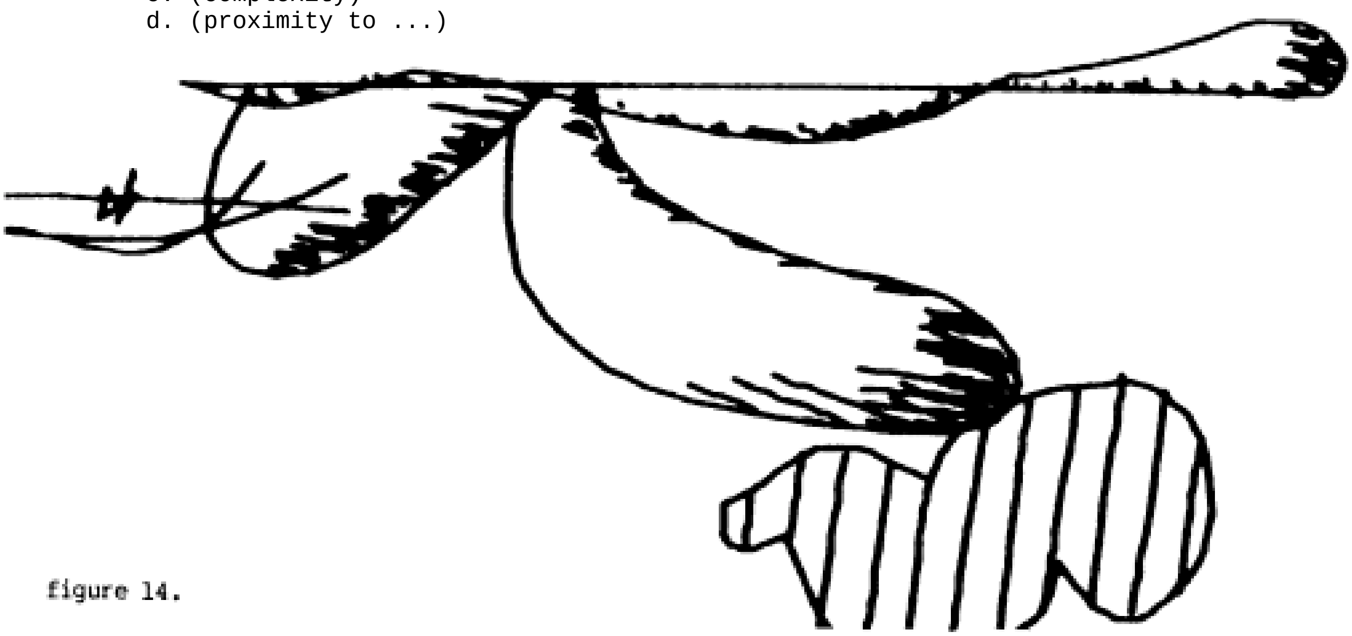


figure 14.

figure 15.

This is a prototype for an expanding class of productions, each responding to a different combination of properties in the first development. Similarly, continuation will require...

If (this is a third development
and the first was a closed form
... properties...
and the second was a closed form
.. . properties. . .)

Then
shade the entire figure:
specification 1:
a boulder with a hole in it
or
specification 2:
a flat shape with a hole
or
specification 3:
a penumbra.

If (this is a third development
and the first was closed
and the second was a series of
parallel lines inside it ...
and the remaining inside space is at
least s...)

Then
do another series of lines:
specification 1:
perpendicular to first...
or
specification 2:
alongside the first...
or
specification 3:
do a closed form in available
space..

(note 7).

The Relationship of Closed Forms and Open Forms.

The same proliferation of options occurs for open-lined structures also, but not to the same degree. One of the interesting things to come out of this program is the fact that open-line structures appear to function quite differently when they are alone in an image than when they appear in the presence of closed forms. There seems to be no doubt that closed forms exert a special authority in an image — perhaps because they appear to refer to objects — and in their presence open-lined structures which in other circumstances might exert similar pressure on the viewer are relegated to a sort of spatial connective-tissue function. A similar context-dependency is manifested when material is presented inside a closed form (fig 15): it is "adopted", and becomes either a detail of the form, or markings upon it. This seems to depend upon particular configurational issues, and especially the scale relationship between the "parent" form and the newly introduced material. This manifestation is important, I believe, in understanding why we are able to recognize as "faces" so wide a range of closed forms with an equally wide range of internal markings following only a very loose distribution specification.

Limits on Development.

At the present time no figure in the program goes beyond three developments, and few go that far, for a number of reasons, in the first place, most of the (formal) behavior patterns in the program were initially intended to model a quite primitive level of cognitive performance, and for most of these a single development is actually adequate. Once a zig-zag line has been generated, repetition, for example —as it is found in existing primitive models — seems limited to those shown in (fig 16).

It has remained quite difficult to come up with new material general enough for the purposes of the program. It is the generality of the protocols which guarantees the generality of the whole, and new material is initiated by the introduction of new protocols. On the level of the procedures which carry out the action parts of the subsequently-developed productions, the approach has been to avoid accumulation of special routines to do special things. There is only one single procedure adapting the protocols of repetition and reversal to the generation of a range of zigzag-like forms, for example (fig 13).

But there has been another, and equally significant reason, for the limitation upon permissible developments. It is the lack of adequate, and adequately important, differentiations in the existing figures. For the primitive model represented by the earlier states of the program it was almost enough to have a set of abilities called up by the most perfunctory consideration of the current state of the drawing: the stress was on the definition of a suitable set of abilities (as represented by the right-hand parts of the productions), and as it turned out it was quite difficult to exercise those abilities without generating moderately interesting results. But for a more sophisticated model it is clearly not enough merely to extend that set of abilities, and the problem of determining why the program should do this rather than that becomes more pressing.

The limitation here can be considered in two ways. One is that I had reached the point of exhausting temporarily my own insights into the image-building process. The other is that I had not made provision in the first versions of the program for being able to recognize the kind of differentiations I would want to deal with — since I could not know at the outset what they were going too be — and thus lacked a structure for developing new insights. This leads to a consideration of my next topic: how the program builds its own representation of what it has done up to any point in the making of the drawing.

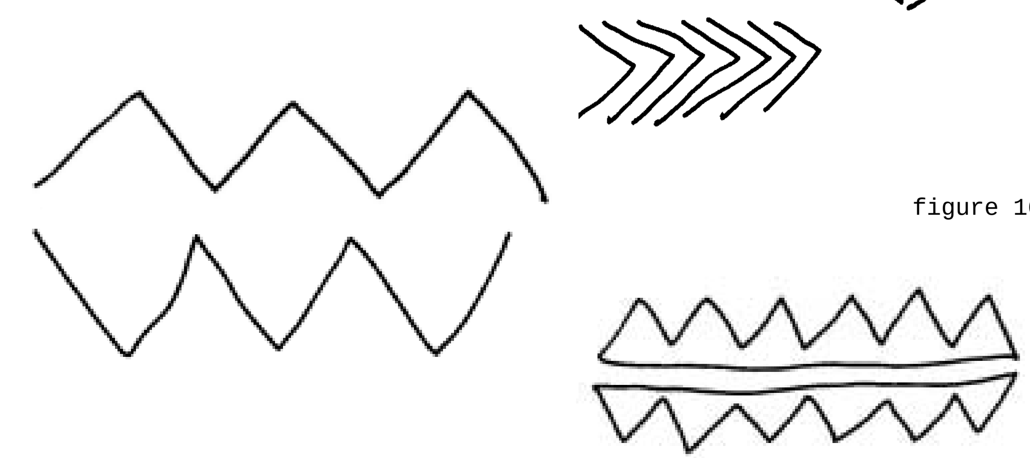


figure 16.

2.4 INTERNAL REPRESENTATION

In the earlier stages of the development of the program, provision had been made for progressive access to the information stored in the data-structure, following the principal that it should not have to access more than it actually needed for the making of any particular decision, in practice, a great deal more was stored than was ever accessed. At the first level of detail the program made use of a quite coarse matrix representation, in each cell of which was stored an identifier for the figure which occupied it, and a number of codes which designated the various events which might have occurred in it: a line belonging to a closed form, a line belonging to an open form, a line junction, an unused space inside a closed figure, and so on. Obviously, it was not possible to record a great deal in this way, and data concerning the connectivity of the figure in particular required a second level of the structure.

This was an unpleasantly elaborate linked-list structure of an orthodox kind. By definition, the kind of drawing AARON makes is not merely a growing, but a continuously-changing, structure. What was a point on a line becomes a node when another line intersects it, and this change has to be recorded by updating the existing structure, which must now ideally show the four paths connecting this node to four adjacent nodes.

Both updating this structure and accessing the information contained within it proved to be quite tiresome, and the scheme was never general enough to admit of further development. As a result, it was used less and less, and decision -making has been based almost exclusively on the information contained in the matrix on the one hand, and in a third level of the structure, a simple property-list attaching to each figure, on the other. The most surprising thing about this simplistic and distinctly ad-hoc scheme is that it was actually quite adequate to the needs of the program.

Explicit Data and Implicit Data.

Human beings presumably get first-order information about a picture by looking at the picture. I have always found it quite frustrating that the program could not do the same thing: not because it made any difference to the program, but because it made it difficult for me to think about the kind of issues I believed to be significant. Part of the problem of using a linked-list structure to represent the connectivity of a figure, for example, derived from the fact that connectivity had to be _explicitly_ recorded as it happened: it would have been much too difficult to traverse a structure of this kind post-hoc in order to _discover_ facts about connectivity. If one could traverse the figure the way the eye does — loosely speaking! — it would not be necessary to give so much attention to recording explicitly all the data in the world without regard for whether it would ever be looked at again.

in short, the primary decision to be made was whether to accept the absolute non-similarity of picture and representation as given, devise a more sophisticated list-structure and drop the matrix representation altogether, or to drop the list-structure and develop the matrix representation to the point where it could be very easily traversed to generate information which was _implicit_ within it. I opted for the latter. A description is included in Appendix 2, though at the time of writing (December '78) the implementation is not yet complete.

2.5 THE FUNCTION OF RANDOMNESS.

This section does not deal with any single part of AARON: randomness is an active decision-making principle throughout the program, and I think it is important to say why that is the case. As a preface, it might be worth recording that beyond the limits of a mathematically sophisticated community most people evidently view randomness in a thoroughly absolutist fashion, and as the opposite to an equally absolute determinism. There is a firmly-held popular belief that a machine either does exactly what it has been programmed to do, or it acts "randomly". The fact that AARON produces non-random drawings, which its programmer has never seen, has given many people a good deal of trouble.

What I mean by "randomness" is the impossibility of predicting the outcome of a choice on the basis of previously-made choices. It follows, of course, that "randomness", in this sense, can never be absolute: if the domain of choice is the set of positive integers, one must be able to predict that the outcome will be a positive integer, not a cow or a color. In AARON the domain of choice is always a great deal more constrained than that, however. The corollary to the notion of randomness as a decision-making principle is the precise delineation of the choice space: in practice, the introduction into the program of a new decision characteristically involves the setting of rather wide limits, which are then gradually brought in until the range is quite small.

Randomness by Design and by Default.

AI researchers in more demonstrably goal-oriented fields of intellectual activity must obviously spend much time and effort in trying to bring to the surface performance rules which the expert must surely have, since he performs so well. I am not in a position to know to what extent "Let's try x" would constitute a powerful rule in other activities: I am convinced that it is a very powerful rule indeed in art-making, and more generally in what we call creative behavior, provided that "x" is a member of a rigorously constrained set.

A number of artists in this century — perhaps more in music than in the visual arts — have deliberately and consciously employed randomizing procedures: tossing coins, rolling dice, disposing the parts of a sculpture by throwing them on the floor, and so on. But this simply derives a strategy from a principle, and examples of both can be found at almost any point in history. It is almost a truism in the trade that great colorists use dirty brushes. Leonardo recommended that the difficulty of starting a new painting on a clean panel — every painter knows how hard that first mark is to make — could be overcome by throwing a dirty sponge at it (note 8). But one suspects that Leonardo got to be pretty good with the sponge! An artist like Rubens would himself only paint the heads and hands in his figure compositions, leaving the clothing to one assistant, the landscape to another, and so on. All the assistants were highly-qualified artists in their own right, however. The process was not unlike the workings of a modern film crew: the delegation of responsibility reduces the director's direct control, and randomizes the implementation of his intentions, while the expertise and commonly-held concerns of the crew provide the limits (note 9).

Randomizing in the Program: Rules and Meta-rules

For the human artist, then, randomizing is not unconstrained, and therefore cannot be characterized by the rule "If you don't know what to do, do anything". Rather, one suspects the existence of a meta-rule which says, "precisely define a space within which any choice will do exactly as well as any other choice". in AARON, the implementation of the low-order rule has the following form:

```
If (a and b and ...n)

  Then p% of the time do (x);
       q% of the time do (y);
       r% of the time do (z);
```

which fills out the description of the format discussed in PLANNING. The same frequency-controlled format is used within the action part of a production in determining specifications:

```
make a closed loop:
specification 1: number of sides
     50% of the time, 2 sides (simple loop)
     32% of the time, 3 sides
        . . .
specification 2: proportion
     50% of the time, between 1:4 and 1:6
     12% of the time, between 3:4 and 7:8
        . . .
specification 3.
```

AARON has only the simplest form of these meta-rules, which are used to determine the bounds of the choice space:

```
if(a) lowbound is La, highbound is Ha
if(b) lowbound is Lb, highbound is Hb
if(n) lowbound is Ln, highbound is Hn
specification taken randomly between
lowbound and higbound
```

where a,b,n are varying conditions in the state of the drawing. No consistent attempt has been made to develop more sophisticated meta-rules. in the final analysis, the existence of such rules implies a judgmental view of the task at hand, and they are consequently beyond the scope of a program like AARON, which is not a learning program and has no idea whether it is doing well or badly.

The Value of Randomness.

What does randomness do for the image-maker? Primarily, I believe its function is to produce proliferation of the decision space without requiring the artist to "invent" constantly. One result of that function is obviously the generation of a much greater number of discreet terminations than would otherwise be possible, and consequently the sense that the rule-set is a great deal more complex than is actually the case. A second result is that the artist faces himself constantly with unfamiliar situations rather than following the sane path unendingly, and is obliged to pay more attention, to work harder to resolve unanticipated juxtapositions. it is a device for enforcing his own heightened participation in the generating process.

This last might seem less important in AARON: the program's attention is absolute, after all. But for the viewer the fact that AARON exercises the function is quite important. There is one level of our transactions with images on which we respond with some astuteness to what is actually there. The fact that AARON literally makes decisions every few microseconds — not binary decisions only, but also concerning quantitative specifications — shows clearly in the continuously changing direction of the line, in every nuance of shape, and succeeds in convincing the viewer that there is, indeed, an intelligent process at work behind the making of the drawings.

****

3. CONCLUSIONS

AARON produces drawings of an evocative kind. It does so without user intervention; without recourse to user-provided data; and without the repertoire of transformational manipulations normal to "computer graphics". It remains now, if not to propose a coherent theory of image-making, at least to pull together those fragments of explanation already given into something resembling a plausible account of why AARON works.

This will be largely a matter of putting things in the right places.

Art-making and Image-making

First: no adequate justification has yet been given for the many references to art and art-making, as opposed to images and image-making, beyond saying that the first are a special case of the second. What makes than special?

Art is a bit like truth. Every culture has, and acts out, the conviction that truth and art exist, no two cultures will necessarily agree about what they are. There is no doubt, for example, that we use the word "art" to denote activities in other cultures quite unlike what our own artists do today, for the quite inadequate reason that those earlier acts have resulted in objects which we choose to regard as art objects. If it is surprisingly difficult to say what art is, it is not only because it is never the same for very long, but also because we evidently have no choice but to say what it is for us.

All the same, no justification is possible for making reference to it without attempting to say — once again! — what it is, and doing so in terms general enough to cover the greatest number of examples. Also, those terms should do something to account for the extraordinary persistence of the idea of art, which transcends all of its many examples.

Briefly, my view is that this persistence stems from a persistent and fundamental aspect of the mind itself. It would be slating the obvious here to propose that the mind may be regarded as a symbol processor of power and flexibility. I will propose, rather, to regard it as devoted primarily to establishing symbolic relationships: to attaching significance to events, and asserting that this stands for that. This is, surely, a large part of what we mean by understanding.

As for art: in its specifically cultural aspects art externalizes specific assertions — the number three stands for the perfection of God, the racing car stands for the spirit of modern man, the swastika stands for the semi-mythical migrations of the Hopi people, or for a number of other things in a number of other cultures. But on a deeper level, art is an elaborate and sophisticated game played around the curious fact that within the mind things can stand for other things. It is almost always characterized by a deep preoccupation with the structures of standing-for-ness, and a fascination with the apparently endless diversity of which those structures are capable. What we see in the museums results from a complex interweaving of the highly individuated and the highly enculturated, and in consequence any single manifestation is bound firmly to the culture within which it was generated: or it is rehabilitated to serve new ends in a new culture. But ultimately, art itself, as opposed to its manifestations, is universal because it is a celebration of the human mind itself.

The Embeddedness of Knowledge

Second: much of what has come out of the writing of AARON has to be regarded simply as extensions to the body of knowledge which the program was intended to externalize. Writing it was not merely a demonstrative undertaking, and it is far from clear what has been raised to the surface and what newly discovered. I have regarded the program as an investigative tool, though for present purposes the distinction is not important.

It remains impossible to give an adequate account of this knowledge other than by reference to the program itself. There are several reasons for this. In the first place, this knowledge does not present itself initially as predominantly prescriptive. The first intuition of its existence comes in the form of an awareness that an issue — closure, repetition, spatial distribution — is significant: the program should be structured in terms of that issue, as well as in terms of all the other issues already contained. In this sense the left parts of the productions might eventually be taken together to represent the set of issues which AARON believes to be worth attending to in the making of an mage. But this stage comes much later, and by this time an individual production functions as part of a fabric of issues, with so many threads tying it to so many knowledge sources, that a one-to-one account of how it achieves its effect is generally out of the question.

In fact, there is only a single example I can call to mind in which an effect can be ascribed with certainty to a single production; a particular class of junction in a meandering horizontal line will infallibly generate strong landscape reference, though only if the

figure 17.

branching at the junction goes off on the lower side of the line (fig 17). This degree of specificity is certainly exceptional, but less powerful as an evacuator rather than more so.

In general, this particular class of junction — it is more easily characterized visually than verbally — tends strongly to denote spatial overlap: but the specific effect is evidently quite context-dependant, and dependant also upon the precise configuration of the junction itself.

"Personality" as a Function of Complexity.

At the higher end of the scale of effects, the problem of saying what causes what becomes more difficult still. I have never been able to understand how there can be such general agreement about the "personality" which AARON's drawings project, or why that "personality" appears to be like my own in a number of respects. Personality has never been an issue on the conscious level of writing code, and I know of nothing in the program to account for it. To put the problem another way, I would not know how to go about changing the program to project a different "personality".

I assume that the personality projected by an image is simply a part of a continuous spectrum of projection, not distinguishable in type from any other part. But I am forced now to the conclusion that these more elusive elements of evocation — personality is only one of them, presumably — are generated out of the complexity of the program as a whole, and not from the action of program parts; that given an adequate level of complexity any program will develop a "personality". This "personality" may be more or less clear in individual cases, and may perhaps depend upon how many people have worked on the program — AARON is almost exclusively my own work — but it will in any case be a function of the program, and outside the willful control of the programmer, if this is the case it seems extremely unlikely that any complete causal account of the workings of a program would ever be possible.

The Continuousness of image-making and Image-reading.

Third: I want to return to the question which lies at the root of this work. What constitutes a minimum condition under which a set of marks will function as an image?

The reader will have noted that much of what has been written here appears to bear as much upon the business of image-reading as it does upon image-making, there is no contradiction: the central issue being addressed is the image-mediated transaction itself, and image-making in particular has no meaningful, or examinable, existence outside of that transaction. Knowledge about image-making is knowledge about image-reading: both rest upon the same cognitive processes. Thus the skilled artist does not need to enquire what the viewer sees in his work: the satisfaction of his own requirements guarantees it a reading in the world, and the explicit individual readings which it will have are irrelevant to him. The trainee artist, the student, on the other hand, frequently responds to his teacher's reading of his work by objecting, "You're not supposed to see it that way", evidently unaware that the reading does not yield to conscious control. Lack of skill in image-making more often than not involves a failure to discern the difference between what is in the image-maker's mind and what he has actually put on the canvas.

It is equally true, I believe, that image-reading has no meaningful existence outside the transactional context: not because the whole event is always present — it almost never is — but because every act of image-reading is initiated by the unspoken assertion "What I see is the result of a willful human act". That is a part of what we mean by the word "image". However much we may amuse ourselves seeing dinosaurs in clouds or dragons in the fireplace, we have no difficulty in differentiating between marks and shapes made by man, and marks and shapes made by nature, and we do not hesitate to assign meaning in the one case where we deny it in the other: unless we belong to a culture with a more animistic attitude to nature than this one has.

In short, I believe that the first requirement of the condition in the question is the undenied assumption of human will (note 10).

The rest of the condition is given by the display of behavior which draws attention to a particular group of cognitive elements, in other words, evidence of cognitive process may be substituted for the results of an act of cognition. An actual desire to communicate — which may include the simple desire to record the appearance of the world — is not a necessary condition.

AARON's strength lies in the fact that it is designed to operate within, and feed into, the transactional context, not to reproduce the aesthetic qualities of existing art objects. It takes full advantage of the viewers' predispositions and does nothing to disabuse them: indeed, it might fairly be judged that some parts :)f the program — the simulation of freehand dynamics, for example — are aimed primarily at sustaining an illusion (note II).

But the illusion can only be sustained fully by satisfying the conditions given above, and once that is accomplished the transactions which its drawings generate are real, not illusory. Like its human counterpart, AARON succeeds in delineating a meaning-space for the viewer, and as in any normal transaction not totally prescribed by prior cultural agreements, the viewer provides plausible meanings.

Standing-for-ness.

Fourthly: there is a multitude of ways in which something can stand for something else, and in adopting the general term "standing-for-ness" I intended for the moment to avoid the excess meanings which cling to words like "symbol", "referent", "metaphor", "sign", and so on: words which abound in art theory and art history. An image, I have said, is something which stands for something else, and of course it is quite plain that I have been discussing only a very small subset of such things.

What are the defining characteristics of this subset?

Before attempting to answer that question, it should be noted that, while AARON's performance is based upon vision-specific cognitive modes (note 12), there are two closely related questions which cannot be asked about AARON at all.

Images of the World and its Objects.

The first of these has to do with the fact that in the real world people make images of things.

How do people decide what marks to make in relation to those things?

It is difficult to avoid the conclusion that image-making as a whole is vision-based, even though it bears directly on the issue of appearances only occasionally. It is my belief that even when an image is not purposively referential — as is the case with AARON — or when the artist seeks to refer to some element of experience which has no visual counterpart, it is his ability to echo the structure of visual experience which gives the image its plausibility (note 13).

The Persistence of Motifs

The second question has to do with the fact that actual image elements, motifs, have been used over and over again throughout human history, appearing in totally disconnected cultural settings, and bearing quite different

Figure 18.

meanings as they do so. what is it that makes the zigzag, the cross, the swastika, squares, triangles, spirals, mandalas, parallel lines, combs (fig 18), ubiquitous, so desirable as imagistic raw material?

My own answer to this question is that the cognitive modes and their dependant behavioral protocols are <u>absolutely</u> ubiquitous, and that the recurring appearance of these motifs is

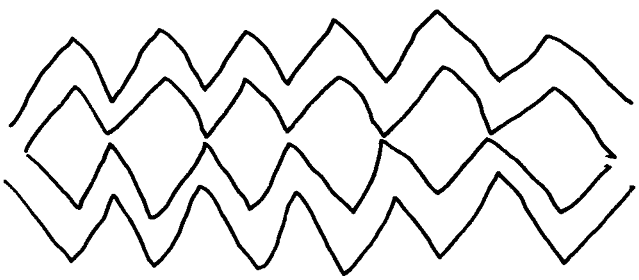


figure 19.

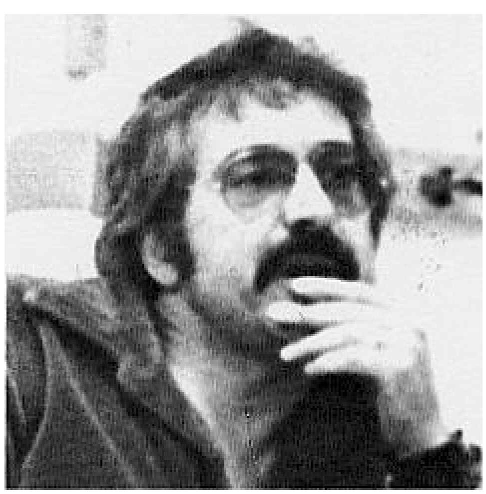
hardly even surprising (note 14). In fact, we have only to start cataloguing the motifs to realize that most of them are simply formed through the combination of simple procedures. The swastika, for example, is both cross and zigzag, just as the mandala is cross and closed form, and the so-called diamond-backed rattlesnake motif of the Californian Indians is a symmetrically repeated zigzag (fig 19).

Taken together, these two questions point to the dualistic nature of image-making. If, as I believe to be the case, it can be shown that the representation of the world and its objects by means of images follows the same cognition-bound procedures as the simpler images I have been discussing, then it will be clear that the form of an image is a function both of what is presented to the eye <u>and</u> of the possession of appropriate modes.

<u>Representation.</u>

I said at the outset that my conclusions would bear upon the nature of visual representation, as distinct from what the AI/Cognitive Science community means by the word "representation". It is still the case that my specific concerns are with what people do when they make marks on flat surfaces to represent what they see, or think they see, in the world. All the same, some speculation is justified about possible correspondences between the two uses of the word.

It is important, for example, to note that the lines which the artist draws to represent the outline of an object do not actually correspond to its cadges, in the sense that an edge-finding algorithm will replace an abrupt tonal discontinuity with a line. In fact, the edges of an object in the real world are almost never delineated by an unbroken string of abrupt tonal discontinuities. If the artist is unperturbed by the disappearance of the edge, it is likely to be because he isn't using that edge, rather than because he has some efficient algorithm for filling in the gaps. Similarly, most of the objects in the world are occluded by other objects, yet it would not normally occur to the artist that the shape of a face is the part left visible by an occluding hand (fig 20).



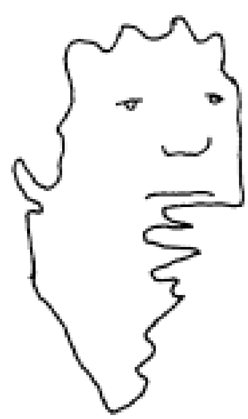


figure 20.

The face evidently exists for him as a cognitive unit, and will be recorded by means of whatever strategies are appropriate and available for the representation (note 15).

It is as true to your meaning of "representation" as to mine, not only that it rests upon the possession of appropriate and available strategies, but also that new strategies may be developed to fit particular concerns. Both are bound by entity-specific considerations, however: considerations, that is to say, which are independent of the particular event or object being represented and take their form from the underlying structures of the entity — the artist's cognitive modes on the one hand and the structural integrity of a computer program on the other.

What is a Representation "Like"?

It could not be seriously maintained that a computer program is "like" a human being in a general sense, and it should not be necessary to point out that a representation in my meaning of the word is not "like" the thing represented, other than in precisely defined senses of likeness. That may not be quite obvious, however, when we consider the idea that a portrait is "like" the sitter. Even though we may be careful enough to say that the portrait LOOKS like the sitter, or that a musical passage SOUNDS like the rustling of leaves, we tend to stop short of that level of detail at which it becomes clear that the appearance of a painted portrait and the appearance of a person actually have very little in common. A representation may be <u>about</u> appearance, but we never confuse the representation with the reality, no matter how "lifelike" it is. in fact, we might rather believe that all representations of a given class are more like each other than any of them is like the thing represented. Life follows its laws, representations follow theirs.

What is an Image?

The <u>purpose</u> of an act of representation is to draw attention to some particular aspect of the represented object, to differentiate that aspect from its context, not to reconstitute the object itself. To that degree we might regard a visual representation as constituting a partial theory of that object and its existence, just as we might regard a computer program as constituting a theory of the process it models. But neither the artist nor the program designer has any choice but to proceed in terms of the modes which are available or which they are capable of developing. In the case of the visual representation, the making of an image, I have tried to demonstrate the cognitive bases of those modes, and also, through my own program AARON, to demonstrate their raw power in the image-mediated transaction.

That, finally, defines my use of the word "image". An image is a reference to some aspect of the world which contains <u>within its own structure</u> and <u>in terms of its own structure</u> a reference to the act of cognition which generated it. It must say, not that the world is like this, but that it was <u>recognized</u> to have been like this by the image-maker, who leaves behind this record: not of the world, but of the act.

APPENDIX I THE TURTLE SYSTEM.

When the real turtle is not running, the program simulates its path, and calculates where it would have been in an error-free world after completing each command. In this case it substitutes a chord for the arc which the real turtle would have traced out. (The straight line segments which may just be visible in the illustrations here are due to the fact that they were photographed off the Tektronix 4014 display, not from an actual turtle drawing.)

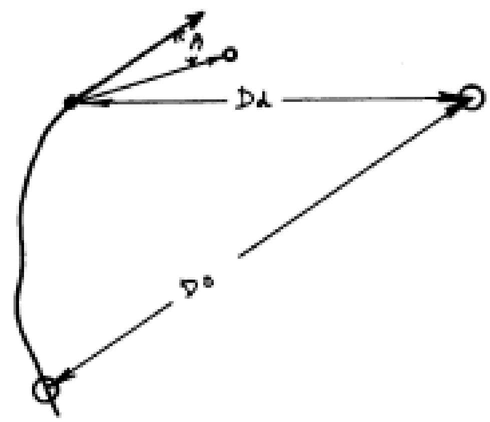
The Navigation System.

The navigation system is correct to about .2 inches: that is an absolute determined by the sonar operating frequency — about 40KHz — and does not change with the size of the drawing. Even with so coarse a resolution the feedback operation is efficient enough for the turtle to do everything on the floor that the program can do on the screen; indeed, if the turtle is picked up while it is drawing and put down in the wrong place it is able to find its way back to the right place and facing the correct direction.

The Dynamics of Freehand Drawing.

There are several complexities in this part of the program which are worth mentioning. One of them is that the program has to be able to accomplish dramatic shifts in scale in the drawing, to make small things which look like small examples of big things: smoothly-curved closed forms should not turn into polygons as they get smaller. This is required both on the issue of shifts in information density and also to maintain implied semantic relationships between forms.

A second complexity is that the movement of the line should convincingly reflect the dynamics of a freehand drawn line, and this should mean, roughly, that the "speed" of a line should be inversely related to the rate of change of curvature: the pen should be able to move further on a single command if it's path is not curving too radically. (The converse of this is that the amount of information needed to specify an arbitrary line should be a function of its rate of change of direction, with the straight line, specified by its two end points, as the limiting case.)



Movement Scaling.

Third, the pen should proceed more "carefully" when it is close to some final, critical position than when it has relatively far to go and plenty of time left to correct for carelessness. This, too, implies a scaling of movement in relation to the state of the local task. Finally, there is the practical problem that for any given number of cycles of a stepping pattern, the actual distance traversed by the pen will vary with the ratio of the turtle's two wheel speeds. Unfortunately, this relationship is not linear, and neither does it provide a useful simulation of freehand dynamics.

Briefly, the line-generating procedure concludes that, given the present position and direction of travel of the pen in relation to the current signpost and to the final destination, it will be appropriate to drive the two wheels at stepping rates *rl* and *r2*, taking *n* steps on the faster of the two. in doing so it takes account of all of the above considerations. The ratio determined for the two speeds is a function of two variables; the angle *A* between the current direction and the direction to the current signpost, and a scaling factor given by the remaining distance *Dd* to the final destination as a proportion of the original distance *Do* (fig II) . This speed ratio then becomes one of the two variables in a function which yields the number of steps to be taken — the distance to be traveled — by the fast wheel: the other variable being the relative size of the block of space allocated to the current figure.

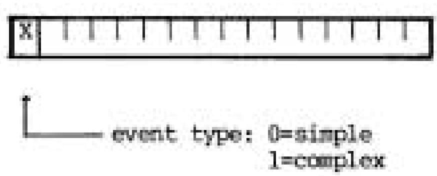
These functions have to be tuned with some care to be sure that each variable is correctly weighted, and to compensate for the turn-distance ratio of the turtle geometry itself. But none of this — or any other part of the program — involves any significant mathematical precision. There are only fifteen stepping rates available, symmetrically disposed between fast forward and fast reverse. The whole program, including extensive trigonometric operations, uses integer arithmetic — this for historical reasons as well as limitations of available hardware — and the geometry of the current turtle determines that it can only change direction in increments of about one sixth of a degree. (The turtle was not until recently interrupt-driven, and for design reasons this incremental direction-change factor was one degree in the earlier version.) Everything relies upon the feedback mode of operation to provide correction and to prevent error accumulation. The point is that a good car driver can drive a car with sloppy steering as well as a car with tight steering up to the point where feedback correction cannot be applied fast enough.

APPENDIX II — MATRIX REPRESENTATION.

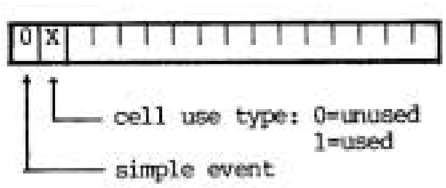
This description is given here primarily because it offers *some* insight into the kinds of considerations which the program believes to be important, and the way in which these considerations are accessed: not because there is anything particularly original from a data-structure point of view.

Much of the detail of the implementation is demanded by the word-length of the machine, and would go away in a larger machine. The intent is to make all the information relating to a particular part of the drawing effectively reside in a particular cell.

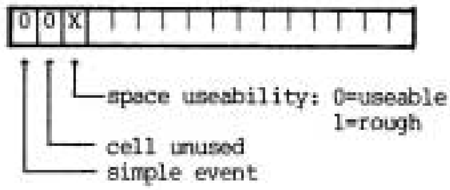
The program uses the single words representing matrix cells in different ways according to what is happening in the cel1s:-



A "simple" event means, essentially, that all the data will be contained within this one word, although it will be seen that its simplicity relates to its use in a more meaningful sense:-

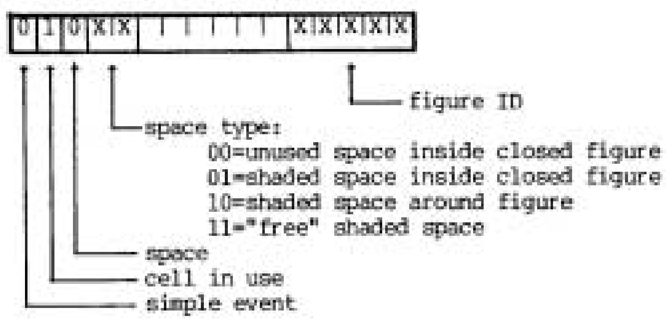


Before beginning work on the drawing, the program "roughens" the surface: that is, it declares some parts to be unuseable for the allocation of space to a new figure, although a developing figure may go into this "rough" space. This is done in order to maximize the rate of change of density across the image:-

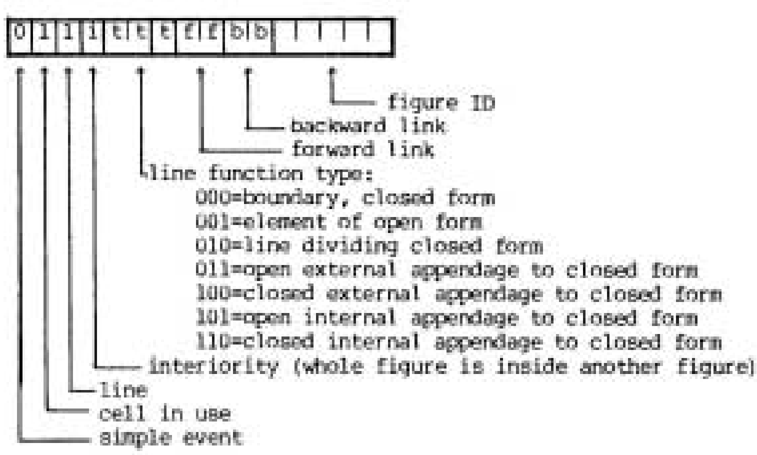


"use" may involve either a line or some special spatial designations:-

in either case, the cell will now have a figure identifier associated with it. The new version of the program uses less figures than the earlier one, and develops than further: a maximum of 32 figures is permitted:-

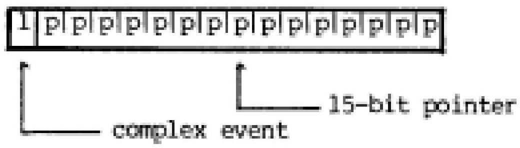


If the cell contains a line, then it can be dealt with as a simple event provided that it is not a line junction of a special kind. In this case the entry designates a line function type;-

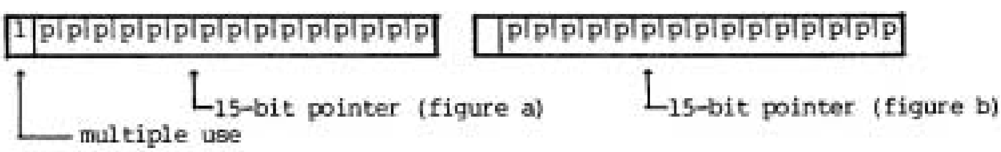


Cell Linking.

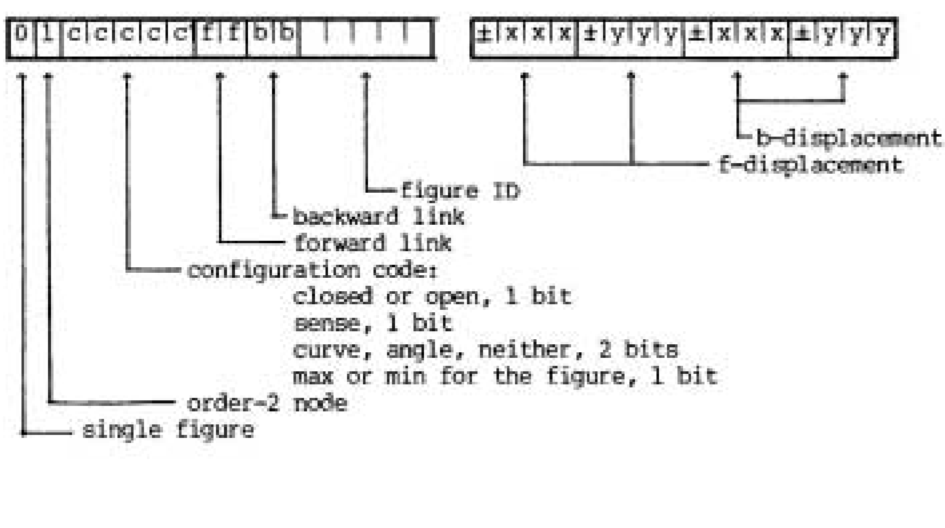
The forward and backward links are a very important device here. Lines are mapped onto the matrix as they are drawn, using an adapted form of Bresham's Algorithm to ensure that strings of cells never include corner-to-corner contiguity, This also means that for any given cell, the line it contains must have entered it from, and will subsequently leave it into, only one of four cells: thus the four-bit linking permits a complete traversal of any series of line segments not involving a complex event.



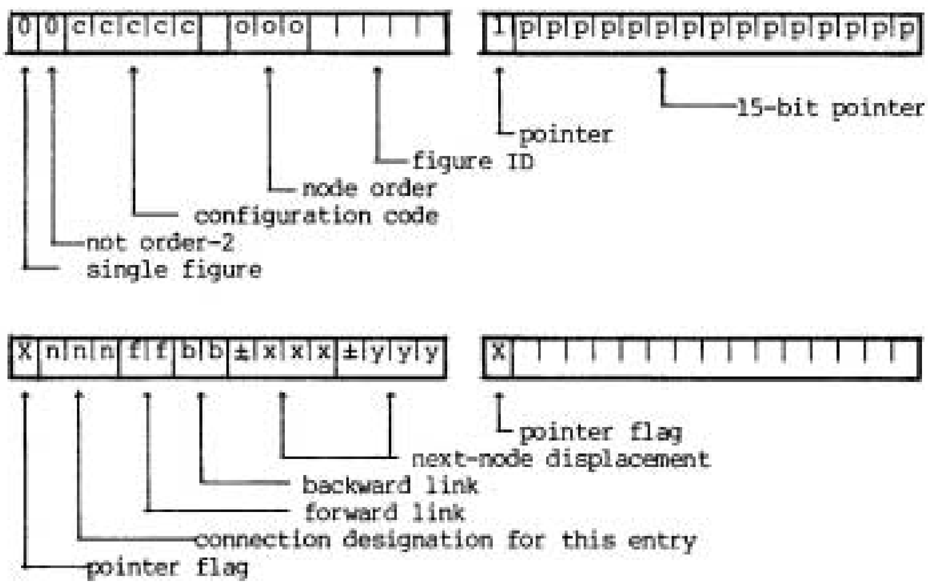
At this point, the single word is inadequate, and it is used as a pointer, words now being allocated in pairs from a freelist. Here again, one level down from the matrix, the words will be variously decoded. In particular, in the event that the cell is occupied by two figures, the two words are each used as pointers to new pairs of words, one for each figure:-



A cell at this level may contain complex events from one or both of two classes: connective and configurational. Configurational events frequently involve order-2 nodes — nodes, that is, which fall on a continuous line — and include sharp angles, strong curvature, and so on. In practice, the program forces complex events so that they always occur within an 8-cell displacement in x and y from another cell, and the location of the next event can then be recorded rather cheaply:-



"Sense", here, means convex or concave if the line is the boundary of a closed figure, and up/right or down/left if it is not. If this is a figure node of any order other than two, one entry will be needed for each adjacent node:-



in addition to the displacements which chain this node to each of its connected nodes. This means that the traversal of the figure as it is represented by the matrix can continue from this point until the next node is reached.

Thus, the entire structure is contained essentially within the matrix, and the short lists which may be tacked onto any single cell serve merely to extend the effective capacity of that cell.

Ideally, this matrix should be as fine as possible: since the resolution of high-grade video is only 1024x1024, a matrix of this size would obviously constitute an extremely good representation. However, there are two considerations which make so fine a grain unnecessary. The first is that the program keeps a full list of all the actual coordinate pairs for each figure as it is drawing it, and can access it should some very precise intersection be required. The second is that the program is designed to simulate freehand drawing, not to do mechanical drawings, and once a figure is completed some approximation to it for purposes of avoidance or even intersection is unobjectionable. The maximum error induced by assuming a point to be at the center of a cell in a matrix of 90x160 will be about 7/8th of an inch in a sixteen-foot drawing: only three times the thickness of the line.

NOTES ON THE TEXT

note 1. The word "representation" is used here in a more general sense than it now carries within the A.I. community: the problem of formulating an internal (machine) representation of some set of knowledge differs from the more general problem primarily in its technological aspects.

note 2. "The Art of Artificial Intelligence: 1, Teams and Case Studies of Knowledge Engineering," Ed Feigenbaum, Proceedings of IJCA15, 1977; pp.1014-1029.

note 3. In the decade before I became involved in my present concerns my work was exhibited at all of the most serious international shows, and I represented my country at many of them, including the Venice Biennale; as well as in some fifty one-man shows in London, New York and other major cities.

note 4. Different from each other, loosely speaking, in the way one might expect a human artist's drawings to differ one from another over a short period of time.

note 5. Written in “C”, under the UNIX operating system.

note 6. I am referring here to differentiations performed in relation to the image, not in relation to the real world, with which the program has had no visual contact.

note 7. The program does not attach semantic descriptors to the things it draws: the terms "penumbra", "boulder" and so on are my own descriptions, and are used here for the sake of simplicity.

note 8. Significantly, from the point of view of my argument here, the dirty marks were intended to "suggest" the elements of a composition.

note 9. The one unconstrained randomizing agent in this scenario, the final cutting of the film by the producer rather than the director, has also demonstrated itself too be devastatingly non-creative.

note 10. "Undenied" is stressed here because there exists an odd case in which the will of the artist is to produce objects which demand the contemplation of their own qualities for their own sake — what they are rather than what they stand for — and which thus seek to deny the viewer his normal assumptions. To the degree that this aim can actually be achieved the resulting object could not properly be called an image, and I doubt whether aesthetic contemplation could properly be called reading. Thus much of XXth Century abstract art falls outside this discussion.

note II. It is worth noting, though, that AARON did mechanical straight-line shading for about two years — it ran faster that way – and in that time only two people ever remarked on the inconsistency.

note 12. I will leave aside the interesting question of whether there are not more general underlying structures which are common to all physical experience. It is presumably no accident that terms like "repetition", "closure", and others I have used in relation to visual cognition are freely used in relation to music, for example.

note 13. The control of the rate of change of information density across the surface of the image, to which I referred earlier, is the most powerful example I know in this regard. The eye is capable of handling units as small as a speck of dust and as large as the sky, but the processes which drive the eye seem always to adjust some threshold to yield a preferred distribution spanning only a few octaves.

note 14. In fact, the more theatrical explanations which range from world-wide migrations to the influence of extra-terrestrial voyagers are not even necessary.

note 15. He is unlikely to treat the boundary between face and hand as part of the face, but as part of the hand, and may very well indicate the full boundary of the face as if he could actually see it.