ARPA ORDER NO. 189-1

MEMORANDUM
RM-5999-ARPA
SEPTEMBER 1969

THE GRAIL PROJECT: AN EXPERIMENT IN MAN-MACHINE COMMUNICATIONS

T .O. Ellis, J. F. Heafner and W. L. Sibley

PREPARED FOR:
ADVANCED RESEARCH PROJECTS AGENCY

The RAND Corporation
SANTA MONICA • CALIFORNIA

THIS DOCUMENT HAS BEEN APPROVED FOR PUBLIC RELEASE AND SALE; ITS DISTRIBUTION IS UNLIMITED.

93

ARPA ORDER NO. 189-1

MEMORANDUM
RM-5999-ARPA
SEPTEMBER 1969

THE GRAIL PROJECT: AN EXPERIMENT IN MAN-MACHINE COMMUNICATIONS

T. O. Ellis, J. F. Heafner and W. L. Sibley

This research is supported by the Advanced Research Projects Agency under Contract No. DAHC15 67 C 0141. Views or conclusions contained in this study should not be interpreted as representing the official opinion or policy of ARPA.

DISTRIBUTION STATEMENT
This document has been approved for public release and sale; its distribution is unlimited.

The RAND Corporation
1700 MAIN ST • SANTA MONICA • CALIFORNIA • 90406

This study is presented as a competent treatment of the subject, worthy of publication. The Rand Corporation vouches for the quality of the research, without necessarily endorsing the opinions and conclusions of the authors.

Published by The RAND Corporation

PREFACE

This Memorandum is the first of a three-part[†] final report on the GRAIL (GRAphical Input Language) Project sponsored by the Advanced Research Projects Agency of the Department of Defense. The study was an integral part of both RAND's and the client's overall program to explore man-machine communications. The Memorandum states the philosophy adhered to in this computer-graphics experiment, describes the problems that were identified, and notes further research directed to these problems. The techniques developed are described in the context of the experiment itself although they are suitable for a larger class of applications.

[†]See also by the same authors: *The GRAIL Language and Operations*, The RAND Corporation, RM-6001-ARPA, September 1969; *The GRAIL System Implementation*, The RAND Corporation, RM-6002-ARPA, September 1969.

SUMMARY

This Memorandum describes the establishment of a common working surface--a cathode-ray-tube (CRT) display--for man and machine. The flexibility of the output channel (CRT) suggests that the input channel should have the same freedom, allowing the man to express himself directly and naturally on a two-dimensional working surface. The GRAIL (GRAphical Input Language) Project explores ways of providing a man with powerful but natural techniques to deal with his problem directly rather than through an agent (e.g., a computer). Experimental software has been developed for the real-time interpretation, evaluation, and reaction to the man's gestures, using a CRT display and a RAND Tablet. The design objectives were: 1) to use only the CRT and the tablet to interpret stylus movement in real-time; 2) to make the operations apparent; 3) to make the system responsive; and 4) to make it complete as a problemsolving aid.

The project deals with the problem of computer programming using flowcharts as a starting point from which to investigate man-machine communications within the above principles. Operations are described that allow the man to specify, edit, validate, document, and exercise his problem description by drawing and gesturing (freehand and in-place) those symbols, characters, and other means of problem expression that he may need. Continuous responses on the CRT display are necessary to minimize distraction and to allow the man to feel that he is dealing directly with the expression of his problem. The various responses (viz., stylus-inking feedback, display-intensity variations, pictorial modifications, and English-language statements) and their appropriateness are discussed.

The underlying consideration in implementation was to shield the man from purely system functions while providing

good response times. Limited computational power and storage led to such real-time data-processing problems as separating the man's problem representation into its pictorial, logical, and positional forms to invoke only the necessary consequences of a stylus action on these forms. The Memorandum examines the solution to some real-time processing problems in terms of dynamic data access and multi-programming.

The present costs of providing adequate feedback are high, although some inroads are being made to alleviate the central-processor load; e.g., the stylus data were originally time-sampled, but they can now be reported to the central processor as differential motion corresponding to the necessary ink feedback. Further, the man's actions can be anticipated by reporting stylus movement even before its switch is closed. In the area of displays, buffered video techniques are available at a lower cost and are less sensitive to information content than random deflection CRTs. Other investigators are developing hardware for picture clipping and scaling to relieve the central processor.

The project's main goal was to identify the problems and study possible methodology for this form of man-machine communications. After the problems have been identified, they can be addressed from the standpoint of economics, which are outside the purview of this investigation.

ACKNOWLEDGMENTS

The authors would like to thank the many reviewers of this Memorandum for their suggestions as well as the following people for their discussion of this work: G. F. Groner, R. Patrick, J. C. Shaw, and R. Turn.

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS vii

Section

I. INTRODUCTION 1

II. GOALS 3

III. THE APPROACH 5

IV. OPERATIONS 6

V. RESPONSES 8

VI. IMPLEMENTATION CONSIDERATIONS 10

VII. DISCUSSION 12

VIII. CONCLUSION 16

REFERENCES 17

I. INTRODUCTION

Recent technology has produced several versatile devices for man-machine communications, one of which is the high performance cathode-ray-tube (CRT) display that can quickly configure and display such information as text, drawings, or even motion.

The flexibility of the output channel (CRT) suggests that the input channel should have the same freedom, allowing the man to express himself directly and naturally on a two-dimensional working surface.

When The RAND Corporation began to explore the capabilities of CRT displays, the evident mismatch between output potentials and existing input capabilities led to the investigation of two-dimensional input devices. The device that resulted, known as the RAND Tablet (see Ref. 1), consists of a pen-like instrument (stylus) used on a two-dimensional surface (printed circuit tablet), which is coupled to a general-purpose computer.

With an evenly matched input-output capability, the CRT face can become the common working surface for both man and machine.† Given the sufficiently expressive input channel, the man can attend exclusively to the display surface (see Fig. 1).

---

†Additional documentation includes a 16-mm color sound movie that portrays the interactive techniques discussed in this Memorandum. The film exhibits some of the language features described here. Copies are available on loan from The RAND Corporation.

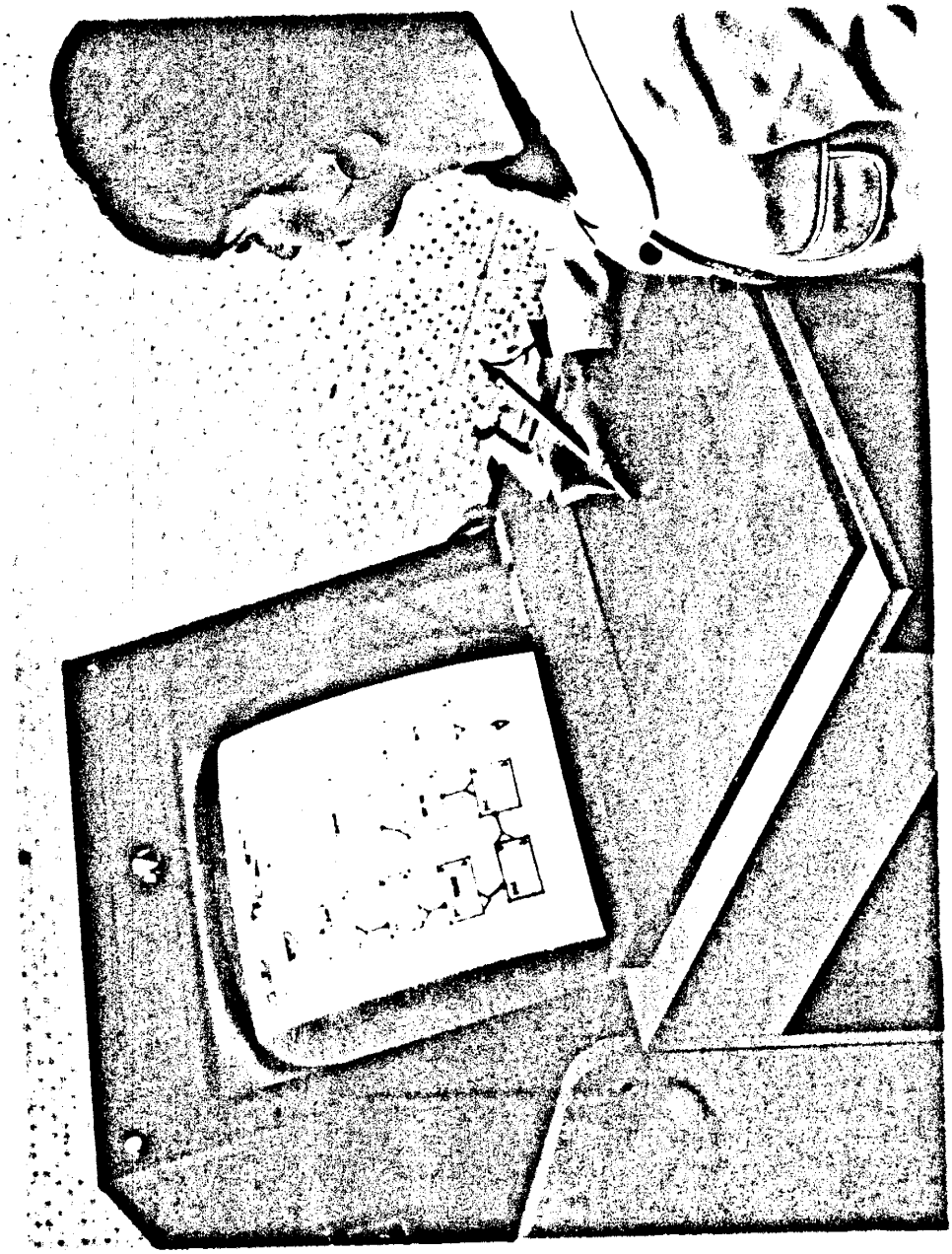


Fig. 1--The GRAIL Console

II. GOALS

The GRAIL (GRAphical Input Language) Project proposed to create an interactive software-hardware system in which the man constructs and manipulates the display contents directly and naturally without the need to instruct an intermediary (the machine); i.e., the display contents should represent, in a very real sense, the man's problem, and allow him to deal directly with it.

For example, consider the construction of a flowchart. An interactive system embodying these features allows a researcher to draw freehand figures and connecting lines; then it immediately replaces these figures with stylized versions of the appropriate size and at the same position to inform him that it understood his actions. If the researcher's actions are in error, the system makes this apparent; e.g., by brightening a symbol or disallowing a connecting line. A conventional system, on the other hand, might require him to specify an object (say by typing R in a prescribed location or by pointing at a displayed figure), its position (by pointing), and its size (perhaps by typing or pointing), then to indicate completion by pushing a button on a keyboard.

The foregoing considerations led to these design goals:

1) Machine-to-man communication to be accomplished solely via the CRT.
2) Man-to-machine communication to be accomplished solely via real-time interpretation of stylus/tablet motions.
3) The environment to minimize ambiguous responses and the operation to be reasonably apparent.
4) The system to be responsive enough for the man to consider the display his working surface with minimal distraction and delay.

5) The system to be complete as a problemsolving aid; i.e., the man should be able to specify, edit, validate (debug), document, and exercise his problem description.

III. THE APPROACH

Computer programming via flowcharts was chosen as a vehicle for the GRAIL project work. Flowcharting is broadly applicable and complex enough to be interesting, as well as being amenable to the proposed communication techniques.

Project specifications demanded a total-system approach to expose thoroughly the problems in developing the techniques for highly interactive systems. Each technique is considered within the context of a real problem in the environment of an operating system; e.g., the real-time recognition of hand-printed symbols--a fundamental part of the project. As part of an operating system, the recognition technique must meet certain real-time and quality standards so it will not degrade the total operating environment. Considered in this role, the merits of recognition features, the structural requirements of the program itself, and the 'contextual' interactions are better understood.

Thus the project followed four parallel paths: 1) definition of operational flowchart symbols; 2) symbol-recognition studies; 3) development of graphical interaction; and 4) planning and implementing the operating system and data organization. The symbol-recognition studies are reported elsewhere [2]. The flowchart symbolism, the graphical interaction, and the operating systems are described in companion Memoranda [3,4].

IV. OPERATIONS

Figure 1 (p. 2) shows the console configuration, which includes a cathode-ray-tube (CRT) display for output and a tablet with stylus for communicating free-hand actions to a computer. The computer controls the information on the display surface except for one spot of light. Conceptually, this spot on the CRT is the pen point and follows the relative pen position on the tablet at all times. A switch in the pen tip senses downward pressure applied to the pen and signals the computer program to accept the pen-track data and interpret it in the context of the information currently displayed.

Thus, the man can describe position, symbols, shapes, or motion direction in place--as if he were actually writing on the surface of the CRT. The man's eyes are always focused on his *working surface*, the CRT. To add or change a symbol in the displayed picture, he draws or prints by applying normal writing pressure to define the separate strokes within the symbol. Immediately (and continuously) the computer presents the path as a visible track on the CRT. This first feedback not only provides the *ink* necessary for the man to draw, but also gives him a real-time indication of exactly what is being interpreted. When this track has described an interpretable symbol, it is replaced (in place) by a normalized symbol. Again, the man can immediately compare his intentions to the computer's interpretation. A major task of the project was to provide this kind of feedback at each of these levels of interpretation.

The directness of interaction attainable depends entirely upon the sophistication of the processes available to interpret relatively natural language elements and relate them to the problem context in real time (man's time). Processes have been developed that allow the man to construct flowchart pictures from a fixed set of recognizable symbols

and to manipulate them. The man may draw these symbols at his own pace in the desired position and size. When the pen is lifted at the end of each handprinted symbol, the computer immediately replaces the free-hand track display with a normalized figure whose type, size, and position are the result of the analysis of that track and its logical implications in the context of the current display. Picture-editing features include the abilities to create, move, stretch, compress, or erase elements of the picture--again, entirely through pen motion.

To minimize ambiguity and distraction, the display remains uncluttered (through constraints applied by the system) and its format consistent from picture to picture. Symbols that the man constructs in a pattern familiar to him remain so, or are changed by the system only when his attention is focused on them; e.g., if the man draws a line representing a control flow from one symbol to another and that flow is unacceptable, the line disappears from the picture immediately. Pictures that are stored and retrieved are not rearranged by the system.

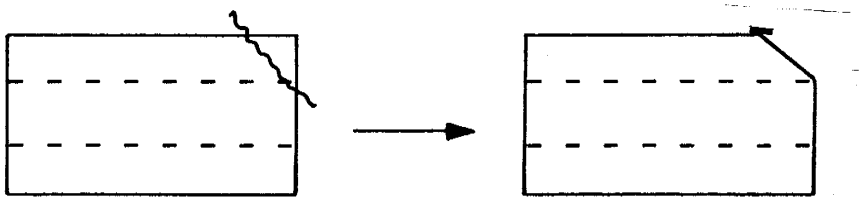
V. RESPONSES

In general, the system responses have been carefully designed to match the man's needs and expectations. The man's most insistent need is to locate the *virtual* position of his stylus, represented by a point of light (direct-feedback point), in the coordinates of the display. Its positioning is a hardware function.

Other responses to stylus input depend upon the context of the pen motion; i.e., the particular picture displayed, the local area addressed on that picture, and, to some extent, the time and direction of the stylus motion. One immediate feedback is an *ink* track--a set of program-generated display vectors that represent the stylus motion on the tablet. The system supplies ink when the man is drawing flowchart symbols and printing characters.

Another response is a change in display intensity. Usually, intensities are at normal brightness; three other intensity levels (off, dim, and bright) are used as responses. The system may apply them locally (where the man's attention is directed) or to the entire picture. The *off mode* indicates that some operation is not applicable in the present environment. The *dim mode* signifies that an operation is in progress; the user cannot affect, for another operation, data and figures thus dimmed. The *bright mode*, applied locally, indicates an error or it focuses the man's attention as the result of his last action.

Still another response is pictorial modification. Some actions may cause the entire display to change; e.g., picture swap. Others are local; below, the user has functionally modified the box by 'chopping off' its corner.



Another response is the English-language statement, which is usually displayed at a familiar fixed location but may be displayed where the man's attention is directed at the moment.

In general, when the stylus motion can be interpreted as an editing or constructing action, the system gives that action its highest priority response; e.g., display inking takes precedence over transfer of data between primary and secondary storage. When a motion implies a more extensive response (e.g., picture swapping), the system may ignore any subsequent pen actions until the response has been completed. In the latter case, for the duration of the response the man should not expect the stylus to be effective.

VI. IMPLEMENTATION CONSIDERATIONS

The high-priority response to constructing or editing operations requires a guaranteed portion of the processor time every 30-60 ms. Normally, that time is used to provide the ink response and to process the stylus track for recognition purposes--or, alternatively, for *moving* a figure about the display. In either case a disturbing servo problem results for the man if the machine does not respond in that time interval.

A portion of the stylus supervision resides in primary store to satisfy the timing demands. Moreover, since the stylus data input is asynchronous, the designers deliberately maintained a form of multiprogramming while reducing the system overhead to a minimum.

Neither all the data nor all the supervisory processes could reside in primary store simultaneously, which led to the hierarchical organization of data, the dynamic allocation of storage, the selection of processing routines from secondary store as needed, and a re-emphasis of the need for multiprogramming to shield the man from purely system functions.

Obviously, every pen action could not be analyzed in exhaustive detail; e.g., a line of text could be created or edited a character at a time until the man was satisfied with its contents. Analyzing that line for errors would be wasteful unless it were complete. The problem then is to detect the completeness of a series of related actions; e.g., the end-of-message signal on a keyboard provides such a termination. This end-of-message problem pervades the entire system. The approach adopted was to treat the start of an action in a geometrically unrelated area as the end of the previous action, which implies that the system would momentarily lag behind the man.

The real-time data-processing requirements led to two additional decisions:

1) The CRT face is not considered a window into a larger two-dimensional space; but rather each picture remains separate and distinct from all others, thus relieving the problems of picture clipping and scaling. Lines of text are excepted.
2) A flowchart is dichotomized into its purely logical meaning (via ring structures) and its purely pictorial form (via display data), allowing some picture modifications to occur without ring manipulations. (An exception is the pages of text. There was no need to distill any logic from the text; hence, its only representation is in a display-list form.)

Of course, the fixed-picture approach does not allow the man to view two or more related pictures simultaneously; i.e., to spread his work out on his desk. A partial solution to this problem is overlaying pictures or splitting the screen between pictures, as appropriate. The display-hardware frame rate introduces enough flicker to limit the usefulness of this approach.

VII. DISCUSSION

The following general remarks should be kept in mind throughout the discussion:

1) The goals presented at the beginning of this paper were subject to the usual constraints of space and computing power. The system was implemented on an IBM System/360 Model 40-G with two 2311 disk drives as secondary store (Fig. 2). For various reasons--the most important being the close control of computer behavior--the project staff built the basic operating system.
2) The choice of computer programming via flowcharting certainly influenced our ability to attain the goals. This choice simplified the problems of construction and control, but the general approach is more widely applicable (see Ref. 5).

The man's ability to focus his attention exclusively on the display is certainly coupled to his ability to effect his intentions directly in place. The seemingly difficult feat of looking one place while gesturing in another (such as typing or driving a car) is really no problem for the man provided the feedback loop is closed quickly enough to avoid a *rubbery* feeling. The present costs of doing so are high, and the system is overloaded under peak demands. This observation is even more disturbing because the operating system was specifically designed for and carefully tailored to this application. However, several areas are being investigated to distribute the central processor's load more evenly.

Some time-consuming aspects of stylus data handling have been identified and cast in hardware; e.g., the direct feedback point, and the *filling in* of the ink track. Originally the stylus position was time-sampled, and varying velocities

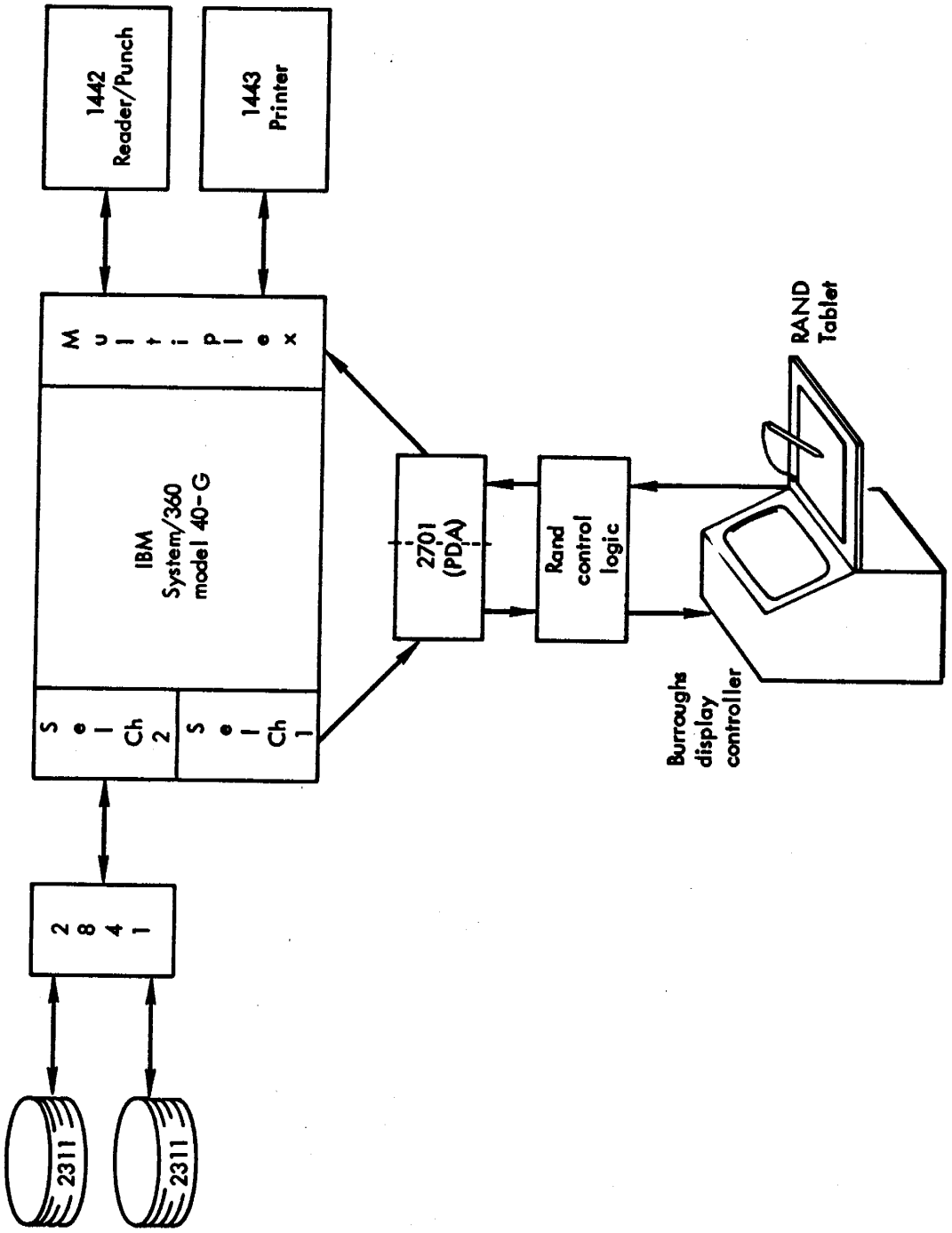


Fig. 2--Hardware Configuration

required different amounts of processor time to provide the inter-sample *ink*. Newer techniques still time-sample the stylus position, but at a higher rate; and the position is reported to the central processor as differential motion (vector directions) directly translatable into ink.

The new stylus-handling techniques provide additional means for smoothing the processor load. Stylus position is available to the computer even while the switch is not closed--in particular, when there is no stylus-motion (as in pointing). By examining its position, the system can anticipate the man as he moves about (or hesitates over) the tablet.

Another aspect of the CRT as a working surface is the amount of information that can be displayed without excessive flicker. A compromise seems necessary to balance the responsiveness of the display against its ability to maintain a steady picture.

Less content-sensitive displays are already available through multi-console buffered video techniques, that amortize the cost of the display-generating equipment over a large variety of simultaneous applications.

In addition, progress is being made by Sutherland and Sproul [6] and others that should bring picture clipping and transforming into reach without adding to the processor load.

The program and data organization--the fact that the system must rearrange its programs or data--should not unduly affect the man. Project experience indicates that relatively large graphic data bases and complex programs present real problems of space and time allocation.

The fast-response requirements coupled with the asynchronous and somewhat unpredictable behavior of the man make priority-driven multiprogramming imperative. No natural place to hide system processing seems to exist.

Finally, the total reliance on real-time stylus data for construction and control raises the question of satisfactory recognition techniques. Inadequate recognition is a source of distraction. The GRAIL techniques were designed to accommodate the project staff; however, experience has proved their wider applicability. The next step will be to apply more contextual information to aid recognition as well as to make the recognizer more adaptable to the man.

VIII. CONCLUSION

The main thrust of the project was to identify the problems and possible methodology of a new form of man-machine communication. Economics were not a prime consideration since they are determined by factors outside the purview of the investigation.

The important result is that such an approach is feasible within reasonable constraints. The future of the techniques lies in extending them to a wider class of problems and reducing their rote aspects to hardware.

REFERENCES

1. Davis, M. R., and T. O. Ellis, *The RAND Tablet: A Man-Machine Graphical Communication Device*, The RAND Corporation, RM-4122-ARPA, August 1964. (Also, *Proceedings of the FJCC, 1964*, p. 325.)
2. Groner, G. F., *Real-Time Recognition of Handprinted Text*, The RAND Corporation, RM-5016-ARPA, October 1966. (Also, *Proceedings of the FJCC, 1966*, p. 591.)
3. Ellis, T. O., J. F. Heafner, and W. L. Sibley, *The GRAIL Language and Operations*, The RAND Corporation, RM-6001-ARPA, September 1969.
4. -----, *The GRAIL System Implementation*, The RAND Corporation, RM-6002-ARPA, September 1969.
5. Sutherland, W., *Semiannual Technical Summary Graphics*, Lincoln Laboratory, Lexington, Mass., November 30, 1967, pp. 24-27.
6. Sutherland, E. I., and R. F. Sproul, "A Clipping Divider," *Proceedings of the FJCC, 1968*, pp. 765-775.
7. Bernstein, M. I., and H. L. Howell, *Hand-Printed Input for On-Line Systems: Final Report for Phase I*, System Development Corporation, TM-(L)-3964/000/00, Santa Monica, California, April 1968.
8. Ellis, T. O., and W. L. Sibley, *On the Development of Equitable Graphic I/O*, The RAND Corporation, P-3415, July 1966.
9. -----, *On the Problem of Directness in Computer Graphics*, The RAND Corporation, P-3697, March 1968.
10. Licklider, J.C.R., and W. E. Clark, "On-Line Man-Computer Communications," *Proceedings of the SJCC, 1962*, p. 113.
11. Sutherland, E. I., "Sketchpad: A Man-Machine Graphical Communication System," *Proceedings of the SJCC, 1963*, p. 329.