HIGHWAY RESEARCH RECORD

Number 455

Application of Interactive Graphics to Transportation Systems Planning

4 reports
prepared for the
52nd Annual Meeting

Subject Areas

22 Highway Design
81 Urban Transportation Administration

HIGHWAY RESEARCH BOARD

DIVISION OF ENGINEERING NATIONAL RESEARCH COUNCIL
NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING

Washington, D.C. 1973

NOTICE

The studies reported herein were not undertaken under the aegis of the National Academy of Sciences or the National Research Council. The papers report research work of the authors that was done at the institutions named by the authors. The papers were offered to the Highway Research Board of the National Research Council for publication and are published here in the interest of the dissemination of information from research, one of the major functions of the Highway Research Board.

Before publication, each paper was reviewed by members of the HRB committee named as its sponsor and accepted as objective, useful, and suitable for publication by the National Research Council. The members of the review committee were chosen for recognized scholarly competence and with due consideration for the balance of disciplines appropriate to the subject concerned.

Responsibility for the publication of these reports rests with the sponsoring committee. However, the opinions and conclusions expressed in the reports are those of the individual authors and not necessarily those of the sponsoring committee, the Highway Research Board, or the National Research Council.

Each report is reviewed and processed according to the procedures established and monitored by the Report Review Committee of the National Academy of Sciences. Distribution of the report is approved by the President of the Academy upon satisfactory completion of the review process.

ISBN 0-309-02188-X
Library of Congress Catalog Card No. 73-15104
Price: $2.00

Highway Research Board publications are available by ordering directly from the Board. They are also obtainable on a regular basis through organizational or individual supporting membership in the Board; members or library subscribers are eligible for substantial discounts. For further information write to the Highway Research Board, National Academy of Sciences, 2101 Constitution Avenue N.W., Washington, D.C. 20418.

CONTENTS

FOREWORD

The papers contained in this RECORD discuss the application of interactive graphics to various aspects of transportation planning and design.

Beilfuss discusses the first phase of a project completed for NCHRP, which consisted of review and definition of an interactive design technique applicable to highway engineering, development of specifications, and the highway design process.

Ruiter and Sussman discuss the potential role of interactive graphics in transportation system analysis. The general characteristics and benefits of interactive graphics are explored, and application in the transportation and planning fields is reviewed.

Gur describes the Interactive Transportation Analysis System, which is a man-computer interactive graphics system designed to serve transportation and urban planners. The author defines the need for interactive graphics in planning, where it might be used effectively, and the major problems in its implementation and use.

Rapp and Gehner describe a man-computer interactive system that was developed and applied to a series of experiments designed to identify the characteristics of high-performance bus rapid transit systems for CBD-bound passengers who reside in a suburban corridor. The solution spaces of each of 12 problem cases derived from the combination of several trip densities, highway network characteristics, and trip-making behaviors are explored in an interactive graphics search process.

INTERACTIVE GRAPHICS IN HIGHWAY ENGINEERING

Charles W. Beilfuss, Control Data Corporation

In September 1970, an NCHRP project was initiated that had as its overall objective the development of an interactive computer graphics software system capable of being used by highway engineers in the design of highways. The first phase of the project was completed in August 1971 and consisted of the following major activities: review and definition of interactive design and roadway perspective system applicable to highway engineering; development of specifications for interactive graphics software, user action requirements, and modifications to current system and programs; review and delineation of interactive graphics terminal hardware specifications; and consolidation of findings and determination of the feasibility of interactive graphics to the highway design process. The end result of the project will be a new method of computer-aided roadway design.

•THE HIGHWAY engineering community was one of the earliest users of, and has continued to be a leader in, the application of computers in engineering work. There has been a continued growth in the sophistication of the computer techniques employed. Even with this continued growth, the engineer, in general, has still not been able to realize the full potential of the computer because of restrictions placed on his ability to communicate with the machine and guide its decision-making process. These restrictions are a direct result of the manner in which information is transferred between man and machine by computer hardware and software. It has been widely acknowledged that benefits to be realized from the computer will be significantly increased when the user can gain more control over the programmed decision process through improved communication techniques.

Most recently, the thoughts of highway engineers are turning to the use of interactive graphics as a means to overcome these communication restrictions. One such thought, toward which briefly positive steps have been taken, resulted in a project to study the feasibility and development of system design specifications for using interactive graphics in roadway design. This project was sponsored by the American Association of State Highway Officials under the auspices of the National Cooperative Highway Research Program. This paper will discuss the way in which this NCHRP project planned to apply interactive graphics to a specific task of highway engineering, final roadway design.

In September 1970, NCHRP contracted with Control Data Corporation to perform phase I of an intended two-phase research project to result in a functioning interactive graphic roadway design system (IGRDS). As of this date, the phase I effort has been successfully completed and the design accepted, but all further funding of this and other similar projects through the NCHRP was terminated by AASHO.

To elaborate a bit more on this work so that the nature of the system can be more readily assessed, I would like to highlight a few key points of this project. The key objectives of the project were

1. To study existing applicable hardware as well as procedures and techniques developed in other interactive graphic applications and
2. To design engineering procedures and software to create a system for interactive graphic roadway design.

The project effort was to emphasize the engineering procedures to be part of such a system. The role of roadway design programs in the system, insofar as the project was concerned, was to be that of a data generator.

The Texas Roadway Design System (RDS) was selected to be used as part of the interactive graphic system to be designed by this project. Selection was based primarily on the broad capabilities of RDS, its modularity, clarity of file design, and the fact that Federal Highway Administration support of RDS would probably make it one that would be much used in the future. The system was not to be machine dependent; it was to be a useful and economically justifiable tool.

Two complete papers would be required to cover both the interactive graphics roadway design procedures and the design recommendations for the specialized interactive graphics software. This paper will be restricted solely to the former concepts and leave the latter for another forum.

IGRDS was to be a research tool that would permit the practical in-service study of interactive graphic techniques by highway engineers. The results of this research study were to provide a practical base for future interactive graphics development in other areas of highway engineering.

The project was made up of work of two different, but related, kinds:

1. Development of procedures for a roadway designer's use of a computer interactive medium and
2. Design and preparation of specifications for computer programs that will allow RDS to work in conjunction with interactive hardware and software in a manner that will permit the performance of the prescribed procedures.

The former category of work consisted primarily of creating new procedural concepts. It involved analyzing each step of the roadway design process and determining what the designer needs and desires to do at each step. As each need was identified, it was necessary to determine of what assistance RDS could be to the engineer at that design step and what interactive graphic techniques could effectively be utilized to provide these design capabilities. The resulting procedures were then recorded in documentary form.

The second type of work included the study of existing interactive graphic hardware and software and the selected RDS to determine what additional interactive graphic program functions were necessary to create a combined system that would effect the desired procedures. This work type also includes determination of whether the proposed interactive engineering procedures were possible. Whenever it was found that the system could not perform in accordance with the procedures first proposed, another approach that was more compatible with the interactive graphic capabilities was sought.

IGRDS DESIGN CAPABILITIES

RDS capabilities are almost exclusively pointed at pure roadway design functions, i.e., design related to horizontal and vertical alignments, earthwork, and geometrics. It is a final design system with limited capabilities for preliminary design and route location because of the way it builds and handles the terrain model. It includes approximately 250 clearly defined programs and subroutines with a large data base of design and terrain-related information, which enables the roadway design engineer to work simultaneously with multiple roadway configurations toward an optimized design. The data base, referred to as project data files (PDF), consists of tabular design information (e.g., template criteria, slope criteria, and equation tables), station-oriented design data (e.g., horizontal and vertical alignments and templates), and individual cross-sectional data (e.g., terrain and design cross sections with related pointers and values). The comprehensive data structure of the PDF permits access by IGRDS to prepare its response to user requests for data and graphic displays.

The design concept of IGRDS acknowledges the probability that design approaches vary among individual designers, and it is to be expected that changes or additions to the system command set will be desired. The design concept is not tied to a particular fixed command set in advance of first trial use; rather, it is one of providing the user with a comprehensive set of commands, software to modify these commands to operate, and a convenient means of modifying the set to fit his own desires.

There are three major divisions of command components that IGRDS must contain to enable the design engineer to function effectively while interacting with the RDS application and its project data files.

1. Geometric, vertical alignment, and cross-sectional design components;
2. Design data input as a direct data entry to PDF; and
3. Analytic displays including mass haul diagrams via earthwork computations, perspective views, and combinations of orthogonal and perspective projections.

The project report contains about 60 design-oriented commands.

The objective of the command structure is to provide a primarily production-oriented system. The concept employed in the user command capability is that, at each step in the roadway design interactive process, the user should have a set number of known, suitable commands available for his selection. Upon selection of any command, the system should act in some specific, expected way, and another set number of suitable commands will be made available for selection in the next step. The selection of each new command is based on the results of the last.

A detailed description of the command structure is not possible in a paper of this scope. Rather, to achieve an understanding of the effect of the commands, the paper will describe the steps an engineer might take as he effects a series of design changes through IGRDS.

DESIGN-ORIENTED COMMANDS

Commands that enable the user to perform design functions in IGRDS occur exclusively in the design command component of the system. They are typified by their effect on the PDF, which is updated to reflect the intended result specified by a selected command.

Horizontal Alignment Geometrics

When working with a horizontal alignment, the user will see it displayed against a grid representing the design coordinate system. There will be no topography or planimetry in the background inasmuch as no such display data are available in the RDS data file; therefore, the design work he performs at the interactive device will generally be already laid out on the design mapping prior to initiation of the interactive process. It is possible to display special grid reference symbols, however, that would represent some controlling feature the user wants to locate in a trial-and-error process.

Upon entry to the horizontal alignment segment of the system, the user defines initial specifications for the display he needs to perform the intended design function. Included in these specifications is a definition of scale and range of interest. Even the largest commercially available display units will not be able to satisfy all users' needs concurrently. If he wants to see the entire station range, the scale will necessarily be very small. This handicap may be overcome, at least partially, by the zoom and window features of the system. He may "zoom in" on a point of immediate interest to a degree where he is able to work with the display. Simultaneously, however, the "window" through which he sees the display permits him to observe an increasingly smaller station range. This impediment may also be overcome by providing him with the ability to move the viewing window from time to time in the design process, or he may "zoom out" to again obtain an overview of the alignment.

The commands will permit the user to build and extend horizontal alignments, or to review them, in a number of ways. The horizontal alignment is reduced to its elemental parts, and each command is aimed at doing something to an element that generally results in a reconfiguration of the alignment. The reconfiguration is reflected in the PDF and displayed on the interactive device. Before the design change is passed on to the RDS application for file update, however, IGRDS will display the new geometric alignment for visual verification.

Upon entry to horizontal alignment, the screen is blank except for the project file identification and the command lists. The user's first action is to choose a number of commands in the display specification list, which result in the painting of the picture shown in Figure 1. The system's simplified space allocation techniques have generally put the labels he requested where they can be read and understood; however, their arrangement may not please him, and he may choose to enter the display administration command list to obtain a more satisfactory presentation.

The original painting produces a clear overwrite conflict between the station equation and the label for PI 4, and the user moves the equation label (Fig. 2a). He also asks that one latitude and one departure on the background grid be identified so that he will be better oriented. Now, for a zoom of PI 2, he moves some other labels, including the north arrow, closer to the alignment so that they will remain in his viewing window.

The zoom action produced three unique pictures in the viewing window, each with an apparent scale double that of the last. Although this final scale does not please the engineer, he is willing to work with it because it is not necessary to use an engineer's scale. The zoom was reasonably good (Fig. 2b). He has an excellent picture of the curve he wants to work with, and he managed to keep the north arrow in the viewing window. The station label for 400+00 also stayed in the window, but it is no longer referencing the correct station tick. Although the graphics are all in the picture, he wants a little more information about them before proceeding with the intended design revision.

Figure 2c resulted from the user's choice of commands from three separate command lists. First, he reentered the display specifications and changed the interval for displaying labels on station ticks. He also asked that a crossroad alignment be included in the display. From the data request table, he obtained curve control stationing, tangent bearings, and complete curve data for the curve at PI 2. Finally, in the display administration command list, he asked for identifying labels on two of the background grid lines and on the main alignment.

Following a rearrangement of the screen display by choosing a number of commands in the display administration command list, the user finally reaches the goal of his entry into the horizontal alignment branch of the system, that of relocating PI 2 (Fig. 2d). He has chosen the command "MVPI↑n" from the design command list and has been advanced in state to accomplish the relocation in the desired manner. At the conclusion of the required actions, the system displays his reconfigured alignment for verification.

Before finally deciding that the design change is what he wants, the user may see the effect of the revision at PI 3. He therefore "windows" down his alignment until the curve at PI 3 is in view (Fig. 2e).

General Geometry

When general geometric computations are performed under IGRDS, an initial display is generated for the user on the display device after he specifies certain parameters that define a skeleton configuration with which he wants to work. These specifications are akin to those occurring in the horizontal alignment segment; the major difference stems from the fact that he has the permanent file of geometrics available to him, and the "skeleton" might be a partially, or fully, completed configuration that he had previously constructed. The relatively small working surface is an ever-present handicap, but it is probably less serious in general geometry. The reach of alignment of interest to him when he is working geometry is much more confined; therefore, he will often be able to see his entire area of interest displayed at a reasonably satisfactory scale. The zoom and window features of the system are available to solve problems that might arise. The fact that he has no background, other than the coordinate grid, is also less serious inasmuch as geometry seldom depends on topography, and special grid reference symbols may be used to represent controlling features of planimetry.

The command set designed to provide geometric design capabilities for IGRDS is based on the existing capabilities of RDS. To be able to perform geometry efficiently on an interactive graphic device requires that the unique hardware and software features of the device be used and the command structure additionally be unique to effect utilization. It is much easier and faster to point to an item of interest than it is to find an element identification and enter it through a keyboard.

A highway-oriented geometric configuration is constructed from unique line elements. Points tie these elements together by mathematical definition. Most of the command set is designed to create unique line elements. The remaining commands enable the user to define relationships between sets of elements. Once specific geometric elements have been defined as a set, the set may be saved in a file for future recall use or modification.

The user enters the general geometry segment with only the command lists appearing on the display. Through display specifications, he gets the centerline of his alignment to appear in a workable orientation, and he adjusts the viewing window to provide space where he needs it. Based on design policies of his highway organization, he establishes the nose of a left-turn lane into an entrance ramp at station 26+35 and then uses commands to establish a construction line normal to the centerline at this station.

With this basic skeleton as a beginning, he proceeds to establish construction points, lines, and curves to build his geometric configuration. Because it takes fewer commands to construct the geometry without controlling the length of line and curve segments, his configuration begins to take on the rather wild appearance shown in Figure 3a. For a time, he is mentally able to keep the maze sorted and continue building the geometric configuration components. The point in time arrives, however, when he must take time out to clear up redundancies before proceeding.

The result of "debris removal" is shown in Figure 3b.

Vertical Alignment Design

The command set for vertical alignment is similar to the horizontal alignment command set. The principal difference between them lies in the manner in which design functions are performed, and this dissimilarity may be traced to the difference in which measurements are made. In horizontal alignment, measurements are made in an infinite number of directions; only two, vertical and horizontal, have meaning to vertical alignment. This simple measuring scheme, together with the use of parabolas with vertical axes to effect curvature, renders the vertical alignment decidedly less complex than the horizontal.

The very things that simplify vertical alignment have somewhat of an opposite effect on IGRDS. The zoom and window features are available in the vertical alignment segment. Windowing is equally as effective as it is elsewhere, but zooming loses some of its effectiveness. A zoom is a scaling device that changes scale equally along both axes. It does not recognize differences in the vertical and horizontal scales. Therefore, a zoom that results in a satisfactory horizontal scale will seldom retain a satisfactory vertical scale, so the vertical scale will have to be reset following the zoom.

The background grid of a vertical alignment will consist of station reference ordinates and elevation reference lines. Design functions may be carried out more effectively in the vertical alignment segment because all the basic background data needed for vertical alignment design may be represented in the display. Most important is the terrain profile at centerline and, if need be, additional terrain profiles along offsets to the centerline. The command set also gives the designer the ability to display a number of "grade control symbols" consisting of points that may be moved horizontally and vertically to positions that control the design of the vertical alignment, e. g., culvert and structure clearance points.

The most effective use of the vertical alignment segment will occur when the design engineer wants to design a profile grade for a new alignment. After bringing out his terrain profile on the display device and setting all known grade controls, he can establish an initial "grade" by describing a single tangent segment extending the length of the specified station range. This might be likened to the initial step of the manual process of laying grade where the profile paper roll containing the plotted terrain profile is spread on a long table and black thread, pinned at one end, is held taut by a dangling weight at the other. Commands to introduce a series of new VPIs replace the manual function. VPIs may then be relocated, just as the designer would move selected pins, to effect an apparent optimization of the grade layout. The manual process is interrupted at this point while the thread is replaced by penciled lines and vertical

Figure 1. Horizontal alignment.

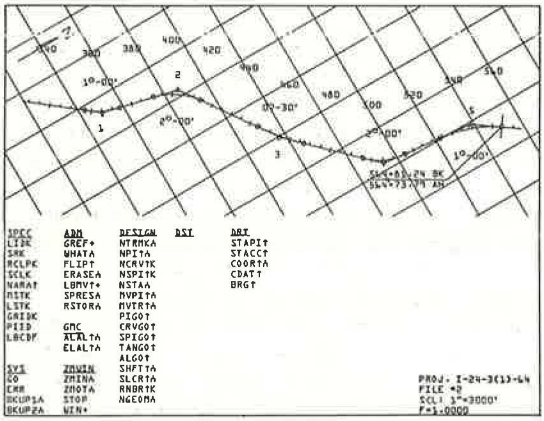


Figure 2. Figure 1 after user-specified commands.

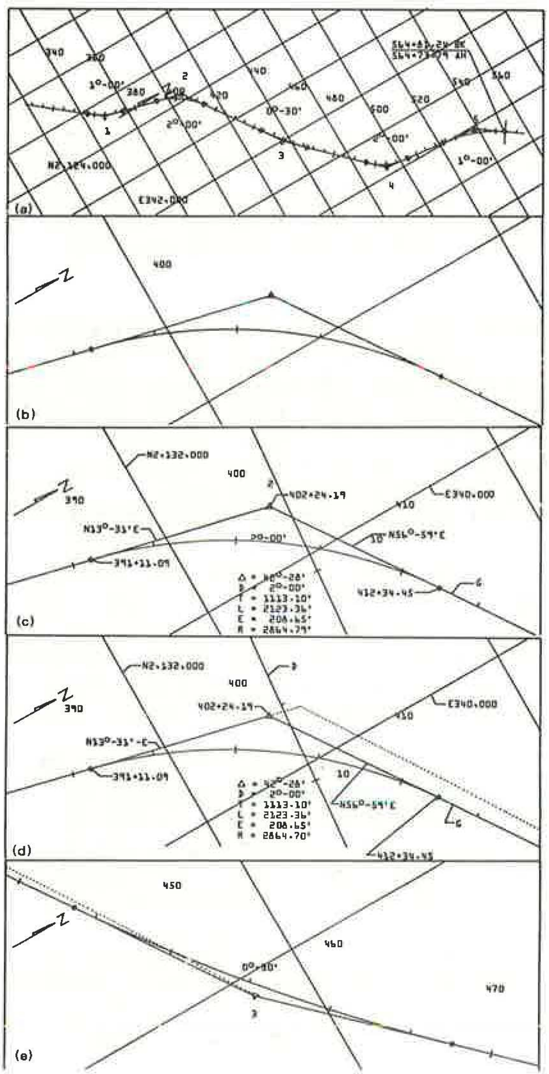


Figure 3. (a) Constructed geometry and (b) geometry with "debris" removed.

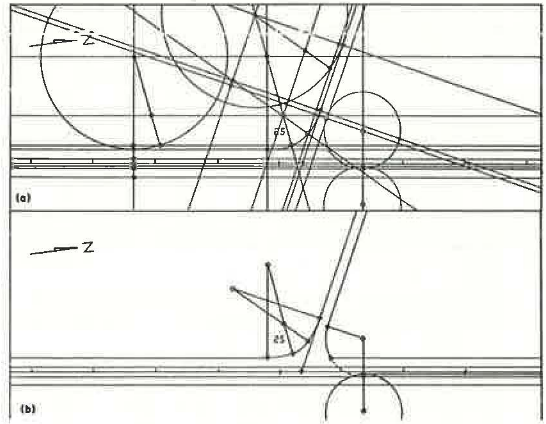


Figure 4. Vertical alignment (a) overview and (b) at working scale.

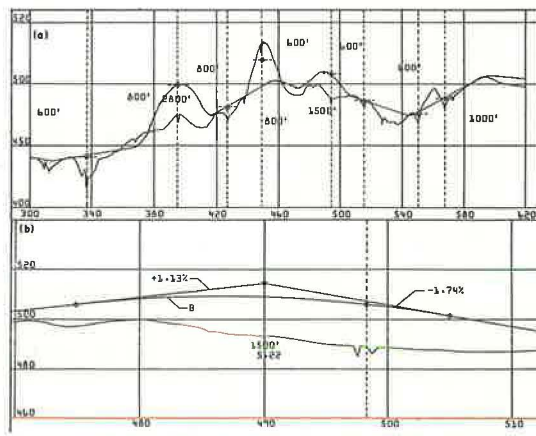
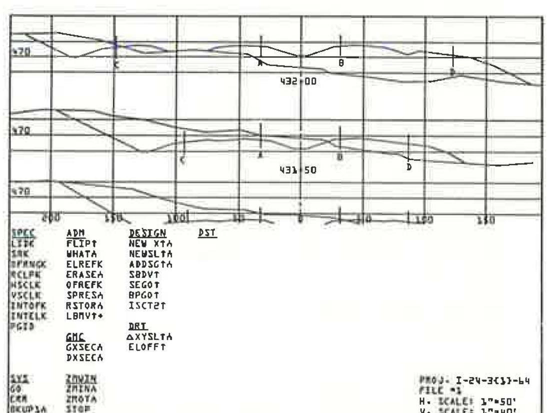


Figure 5. Terrain cross section to be interpolated.



curves, and further adjustments are made without the benefit of the thread. The IGRDS user retains his "thread" throughout the process, and, once vertical curves are defined, they will follow VPI adjustments, which will enable him to see their exact effect on the design.

Approximately 6 miles of vertical alignment and terrain profile at centerline are shown in Figure 4a with vertical curve length labels. Curve control symbols and tangents to the parabolic curves have been suppressed in the display because they are distracting at such a small scale. A number of grade control symbols have been brought out and maneuvered into meaningful positions.

Figure 4b shows the results of a combination of zooming in on Figure 4a through four successive frames on the VPI at station 490+00 and respecifying a vertical scale. The grade control symbol originally specified remains in the picture, and, in addition, the user asks for the parabolic constant to appear with vertical curve length.

Cross-Sectional Design

Four separate segments are included under the cross-sectional design component of IGRDS. Only one may truly be considered a design function. The others, given below, are design aids:

1. Edit terrain cross sections,
2. Interpolate terrain cross sections,
3. Edit and adjust design cross sections, and
4. Trace profile of selected cross-sectional points.

In the first segment, the roadway design engineer can have the digital terrain cross-sectional data residing in the PDF displayed in graphic form for purposes of visual inspection. No interactive data purification capabilities are provided for two important reasons. First, when the display reveals suspicious data, correction should be preceded by verifying that the data are erroneous. There is danger that the ease of interactive purification would introduce as many errors as it corrects. Second, even when an obvious error is found, the user is not likely to have the information at hand that is needed to make the correction, and, even if he does, the time required for the research would detract from the efficiency of the interactive process.

Fully automated interpolation of terrain cross sections has never been a complete success, primarily because there is no programmed fail-safe method of logically finding longitudinal lines of interpolation, i.e., ridge and valley lines. The interaction of IGRDS provides user decision capabilities to assist in this task.

A shallow, natural drainage channel meanders diagonally across the alignment where a terrain cross section is to be interpolated (Fig. 5). Normal longitudinal interpolation would be erroneous, so the user tells the computer how to do it correctly. There is insufficient data on the right side of the two base cross sections to guarantee complete accuracy at the same end of the interpolated section, but the ground trend indicates that any errors that might be introduced would be small.

The interpolation provision in IGRDS will divide the display area on the graphic device into two subareas. One area displays the base cross sections for the interpolation process. The other displays the centerline (normal lines representing the base cross section superimposed with visible offset points representing break points) and a line representing the subject cross section to be interpolated. The user can define all the lines of interpolation and see them represented in the plan view, or he may ask that the RDS application generate these lines, by using its own technique, and that IGRDS display them. He may then purge erroneous or redundant lines and replace them with lines of his own choosing. Finally, he may trigger the interpolation process and have the interpolated cross section displayed for his inspection.

The one true design function of the cross-sectional design component occurs in its third segment. The roadway design engineer can access the design cross sections logically generated by the RDS application and make individual adjustments dictated by engineering judgment.

Figure 6 shows the typical cross-sectional display used for final adjustment of the section. The two roadways of a divided highway and a pair of ramps diverging as sta-

Figure 6. Cross-sectional display used for final adjustment.

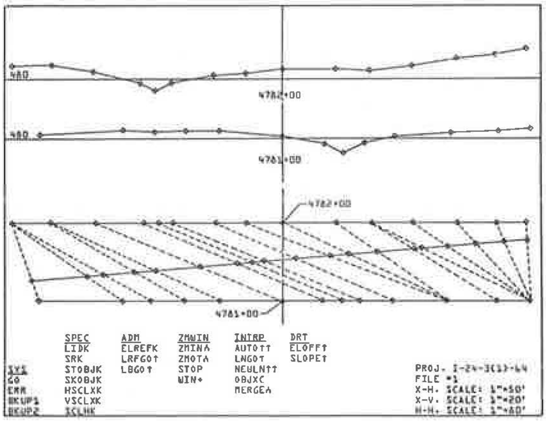


Figure 7. Profile trace of selected point.

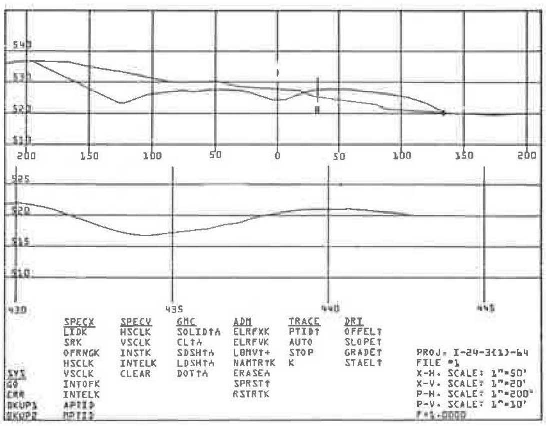
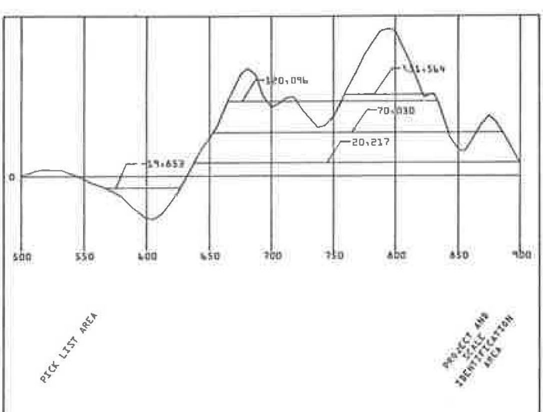


Figure 8. Mass haul diagram.



tioning increases are illustrated. Note that, although data are carried in the display files for the complete cross section at station 431+00, the viewing window is not quite large enough to show all of it. Windowing downward slightly would probably result in a visible display containing three full cross sections.

The final segment of the cross-sectional design component also divides the display surface into two display areas: one reserved for the display of a single terrain-design cross section and the other for a profile view. The user may designate a station range of interest and other pertinent specifications that will result in the display of the first cross section of the range indicated and a background grid of station reference ordinates and elevation reference lines for the profile. He may then choose a point of interest on the displayed cross section and request that a profile be traced that is representative of that point through the station range. When the project data file of the RDS application contains information that makes the chosen point identifiable on each successive cross section (e.g., slope stakes are generally identifiable), the process of obtaining the profile trace may occur automatically. If the point is not internally identifiable, the user must perform the identification on each successive cross section.

When point identification takes place, the sequence of events that follow on the display device provides the engineer with an attractive new design technique. Each successive cross section is displayed sequentially with its profile grade reference point fixed in one spot on the face of the interactive graphic device. Elevation reference lines are adjusted with each new display to accomplish this. With each new cross-sectional display, a line segment is added to the profile trace. When the last cross section in the station range appears, the completed trace is displayed for whatever analytical purpose the engineer had in mind.

One such purpose is to determine the requirements for special ditches and toe ditches. For instance, by tracing the profile of a slope stake of a section of embankment and simultaneously watching the slope trend of the existing ground adjoining the side slope, the engineer can quickly discover small pockets of ponding water where toe ditches are needed for drainage. The command set allows the user to request and obtain precise data from the two views, so that in another design step he may enter the data to include the toe ditch.

In Figure 7 the profile of the left slope stake is traced as successive cross sections appear in sequence on the display screen. The vertical scale in the profile view is purposely warped to make variations in slope more recognizable. Note that, if the sequenced cross section views had shown that the terrain were sloping downward to the left against the referenced slope stake between stations 430+00 and 440+00, a pocket or ponding would be indicated that would need to be drained with a toe ditch.

COMMANDS RELATED TO PRESENTATION OF ANALYTICAL DISPLAYS

In addition to commands that can change the nature of the roadway design itself, there are those that, at the user's request, provide him with displays of the data representing the current status of the roadway.

The first of the analytic displays permits the design engineer to trigger earthwork computations over a specified station range in which specified roadways contribute to the volumes. He describes various display parameters, and the mass diagram resulting from the earthwork computations appears for his analysis. He will need the ability in the command set to strike balance lines and to adjust them on the display. He will also need the ability to request data that the mass diagram is uniquely qualified to give and data resulting from the earthwork computations themselves.

Figure 8 shows a mass diagram with a zero reference ordinate originating at the initial specified station. The user may strike any number of supplementary balance lines and request data to be computed for volumes and haul quantities within each balance and between balance lines.

Perspective drawings, through use of real design and terrain data, of a highway as viewed from a driver's eye position will be possible with IGRDS. This will be accomplished by the perspective programs developed by the U. S. Department of Transportation, Federal Highway Administration, Region 9, that are currently being installed in RDS.

The perspective view component of IGRDS will permit the user to view selected reaches of his alignment when he specifies a station range of interest for each reach, together with a vantage point and a sight point.

The system described in this paper is both a production tool and a research tool. Such a system would be one of the first useful systems to aid the highway engineer through interactive graphic techniques. As such, it would leave much room for future improvement, in both breadth and sophistication. It would, however, provide the means to study these techniques in a production environment and to obtain the reactions and evaluations of practicing design engineers.

INTERACTIVE COMPUTER GRAPHICS IN TRANSPORTATION

Earl R. Ruiter and Joseph M. Sussman, Department of Civil Engineering, Massachusetts Institute of Technology

This paper considers the potential role of interactive graphics in transportation systems analysis. The general characteristics and benefits of interactive graphics systems are explored, and a survey of existing systems is presented. In particular, applications in the transportation and planning fields are reviewed, and conclusions on the use of interactive graphics systems for transportation systems analysis are presented.

•THE FIELD of transportation systems analysis is becoming increasingly complex. The number, size, and scope of the transportation systems the profession is being called on to consider are growing exponentially. Public awareness of the political, social, and economic impacts of these systems is likewise increasing. It is clear that the analyst needs help as he considers the many alternatives open to him in making transportation decisions.

One largely unexplored mechanism for assisting the transportation analyst in his decision-making is interactive computer graphics. The purpose of this paper is to consider the possibilities of interactive computer graphics for use in the transportation field. To accomplish this, the general characteristics and benefits of interactive graphics systems are explored, and a survey of existing production interactive graphics computer systems is presented. A limited benefit model highlighting the important aspects of such systems is introduced.

Then, applications of interactive computer graphics in the transportation and planning fields are reviewed. Based on the above, conclusions on the use of interactive graphics systems in general and in transportation in particular are presented.

BACKGROUND

Interactive computer graphics systems have a very short history. The initial major pioneering efforts took place at M.I.T. and General Motors in the early 1960s. During this period, Ivan Sutherland experimented at M.I.T. with a cathode ray tube (CRT), light pen, function console keyboard, and the experimental TX-2 computer. This effort led to the SKETCHPAD system, with applications in the areas of drafting and structural analysis (25). During the same period, the Design Augmented by Computer (DAC) project was under way at General Motors. This project led to the development of DAC-1, a system used for automotive design (11). Following these initial efforts, interactive graphics systems were developed by industry and universities. A number of these efforts are discussed in this paper.

Although a great deal has been accomplished in the 10 years since Sutherland's original breakthrough, the potential of interactive computer graphics in production use is largely unrealized. There have been a few efforts to utilize interactive graphics in production work, but these have tended to be isolated examples.

The principal reasons for the slow development of interactive computer graphics were

1. Hardware costs—Interactive displays tended to be extremely expensive in terms of the display device itself, support hardware, and the amount of machine time used during operations.
2. Software support—Interactive graphics applications are only meaningful if they are integrated with a well-designed application software system, and few well-designed

application packages existed with which interactive graphics capabilities could be integrated.

3. Man-machine interaction—The application of interactive graphics was not well understood. Analysts with some computer background were oriented toward batch processing computer operations. The introduction of such concepts as remote job entry, time sharing, plotting, engineering input-output stations, and computer graphics radically changed the environment between the analyst and the computer. Significant questions were raised relating to where, how, and to what extent interactive computer graphics fit in the decision-making process.

The past 10 years can be viewed as the decade of research and experimentation in computer graphics. This period corresponds to the decade of 1950-1960 when computers were first introduced. The potential of computers was acknowledged, but the realization of that potential was a difficult, lengthy, and expensive process. A dramatic change occurred between 1950-1960 and 1960-1970. Hardware costs declined, and processing capabilities increased. Applications evolved from simple numerical computations to large-scale information systems. The analyst began using computers as a real tool in decision-making rather than as simply a substitute for the slide rule and desk calculator.

The same kinds of changes are now occurring with regard to interactive computer graphics. Significantly less expensive display devices are available. Application packages that can utilize interactive graphics capabilities are being developed. The analyst is beginning to understand better an interactive graphics environment and how the many capabilities are best used. All in all, these changes point to increased use of interactive computer graphics over the next decade. Given that, it is useful to examine where we have been and what has been accomplished in the field to date. In this light, this paper proceeds to establish (or, more accurately, review) the perceived advantages of interactive graphic systems and then survey existing graphics systems with respect to these advantages.

DEFINITIONS AND CHARACTERISTICS

A useful definition of computer graphics given by Siders (22) is as follows: "The term . . . refers to the concept of man communicating with a computer by means of graphical symbols such as lines, curves, dots, and so forth." However, interactive graphics implies characteristics above and beyond this definition. In this paper, the following considerations pertain: Computer graphics is interactive when man and computer may engage in a dialogue. In particular, interactive computer graphics involves both graphical output and graphical input. Computer graphics is interactive when the graphics capabilities can communicate with an analysis system. ("Analysis system," as used here, refers to the computing capabilities that can exist without computer graphics, e.g., a structural analysis system, but that are enhanced if used with computer graphics.) Through the graphics system, the user has control over the operation of the analysis system, and the user can observe, often dynamically, the progress of the analysis system as it proceeds through its runs.

An interactive graphics system may be interactive in one or the other or both of the senses described. The desired degree of interaction depends on an evaluation of the added costs of highly interactive hardware and software systems versus the added benefits of these systems. These benefits will, of course, vary from problem area to problem area. However, in general, they can be classified in three general catagories: time savings, the development of better alternatives, and cost savings.

Time Savings

A number of authors have quoted various amounts of time savings associated with the use of computer graphics systems. Typically, these quotations vary from a ratio of time with computer graphics to time without computer graphics of anywhere from $\frac{1}{2}$ to over $\frac{1}{1,000}$. This wide range can be narrowed by differentiating between direct time saved and elapsed time saved. Direct time savings can be expressed as the ratio of the time required to perform a specific task entirely with a computer graphics system to the time required to perform the same task entirely without the use of computer graphics. Typical ratios for direct time savings are in the range of $\frac{1}{6}$ to $\frac{1}{1,000}$. Similarly, elapsed time savings can be expressed as the ratio of the total time required to complete a major design effort with the use of computer graphics to the total time required when computer graphics are not used. Typical ratios for elapsed time savings are in the range of $\frac{1}{2}$ to $\frac{1}{8}$.

Development of Better Alternatives

Computer graphics can lead to better solutions to engineering problems for two major reasons. First, the direct time advantages described in the previous paragraph allow the engineer to do more experimentation. In problems that involve trial-and-error analysis, as most engineering problems do, the engineer can explore a wider range of potential solutions. The probability of finding a better solution is therefore much greater when the engineer uses some portion of the time savings available to him to explore his "solution space" more carefully.

Second, the engineer is more likely to obtain an intuitive "feel" for the problem he is observing when he is able to make modifications almost instantaneously and to observe the results of these modifications in a graphical display. Given this "feel," the engineer is more likely to propose better solutions to his problem. This phenomenon has been described by many as "synergy," or the 2 + 2 = 5 effect, in which the capability of the engineer to solve problems heuristically (the trial-and-error approach) is combined via graphics communication to do a better job than either can do separately.

Cost Savings

In most applications, computer graphics will ultimately be evaluated by using monetary measures. In the private firm, the measure is usually the contribution to increased profits. In the case of a public agency concerned with transportation planning, the measure should be net benefit or net cost savings. This type of evaluation, however, must be based on a definition of costs and benefits that is broad enough to include all who will be ultimately affected, both positively and negatively, by a computer graphics system for transportation planning. Some cost savings will be directly felt by the public agency that has at its disposal a transportation graphics system. Others, such as reduced operating costs on highways due to the design of better alternatives, will be less directly felt but no less important in the evaluation of graphics systems.

The advantages discussed in the previous paragraphs can be expressed as cost savings. Direct time savings are reflected in reduced engineering costs. Also, engineers who are able to continue work on a project, without time-out to wait for computer output and drawings of the results, can keep their trains of thought in motion with no delay to recall and rethink previous work toward a solution.

Elapsed time savings reduce the time from project initiation to project completion. In the case of transportation facilities, this normally means that travel cost or time savings or both will begin sooner for the users.

Better alternatives can also result in significant cost savings. Alternatives may be better in that they cost less to construct. And, as mentioned above, alternatives may be better in that they result in lower operating costs.

An additional potential area of cost savings also exists: Computer graphics can reduce the overhead cost of engineering and planning work by reducing the number of hard-copy documents, both graphic and tabular, that must be produced and maintained. In some engineering operations, about one-half of the wages and salaries are devoted to drafting. If the analyst can see, almost instantaneously, the information he needs on a computer graphics display, he can do without many of the hard-copy displays and computer listings that tend to clutter his working space and that cost significant amounts to produce and maintain.

We have discussed the potential advantages of computer graphics without being specific in terms of existing or proposed systems. Also, we have ignored the disadvantages or costs of computer graphics. These considerations are included in the remaining sections of this paper, where specific existing systems are discussed.

EXISTING SYSTEMS IN PRODUCTION USE

The existing interactive graphics systems with engineering applications, for which some information on costs, benefits, or overall cost-effectiveness is available in the literature, are discussed in this section. For the evaluation of the systems to be valid, we felt that they should be working versions rather than academic systems, research efforts, or system proposals. At the present time, the number of such systems is very limited. This is undoubtedly due to the relatively short period of time during which interactive graphics hardware has been generally available, the even shorter time that the required software (plotting packages, communications packages suitable for interactive graphics, and time sharing) has been available, and, until recently, the high costs of both hardware and software.

It was hoped that one or more systems with transportation planning capabilities would be found that met these criteria, but no such systems were found. The major systems meeting the criteria have all been developed and used by industrial concerns.

General Motors

The first major development of an interactive graphics system by an industrial firm was the DAC-1 effort, begun by General Motors in about 1959 and not announced until 1964 (11). The hardware for the system was built by IBM to GM specifications and later became the prototype of the IBM 2250 console. The GM system has been developed to be useful in various portions of automotive design, including body styling, crash simulation, and automatic drafting. Elapsed time savings for the complete process of automobile design of 2 years using DAC-1 versus 4 years using noncomputer graphics procedures have been quoted. Based on the experience with DAC-1 using second-generation computers, the system has been modernized to form DAC-2, a system based on the IBM 360/67 and 360/65 computers and 2250 graphics terminals. DAC-2 is now in normal production use, and present plans call for expanding the system as time goes on.

Lockheed-Georgia Company

The Lockheed-Georgia Company has been among the leaders of the aerospace firms in developing and applying computer graphics (8, 17). Its prototype work was done by using CDC Digigraphics hardware. The production version initially used DEC 340 devices and later IBM 2250 devices. The first working capability was the generation of automated machine tool control tapes. Interactive computer graphics was used, replacing the standard Automatically Programmed Tools (APT) programming language, to prepare the tapes, which control the manufacture and finishing of small parts. This capability was first available in 1965 and has been found to reduce the tape generation elapsed time from a week to 24 hours. In addition to the time savings, major benefits of the capability are that it eliminates the need for programming expertise in the APT language and results in fewer rejected parts due to faulty machine control tapes.

Following this initial success with interactive graphics, Lockheed-Georgia has gone on to add a number of capabilities to the system. Some of these are as follows:

1. Structural analysis of airframe sections,
2. Generation of airplane fuselage surfaces,
3. Design of printed circuit layouts,
4. Simulation of aircraft landings,
5. Interpolation and data smoothing of three-dimensional airframe test data (computer graphics has reduced this task from being a job of 1 or 2 weeks to one of a few minutes), and
6. Placement of parts on large standard-size surfaces.

Mobil Oil Company

Using IBM 2250 hardware, Mobil Oil Company has developed interactive computer graphics capabilities (9) in the following areas:

1. Design of fractionating towers (computer graphics has reduced the elapsed time for this task from months to hours and has resulted in more efficient and cheaper designs),
2. Analysis of seismic data for oil exploration, and
3. Layout of pipeline complexes and control systems.

McDonnell Douglas Corporation

Among the extensive interactive computer graphics systems surveyed, McDonnell Douglas is unique in that it was developed largely by its ultimate users in an open-shop environment (9, 14). This approach has led to a wide range of capabilities:

1. Input and editing of three-dimensional aircraft shapes with direct time savings of 1 month to 10 minutes,
2. Structural analysis,
3. Analysis of airfoil performance with direct time savings of 6 weeks to minutes,
4. Simulation of flight paths with direct time savings of from 4 weeks to 1 hour,
5. Scheduling of projects using PERT,
6. Prediction of passenger seat-miles using an econometric model,
7. Comparison of the costs of surface and air freight in a distribution cost model with direct time savings of days to minutes,
8. Scheduling of airlines,
9. Analysis of airport runways,
10. Continuous system modeling program with graphic output, and
11. Calculation of the return on investments.

This wide range of graphics capabilities includes a number that are of a type foreseen in a transportation graphics system, especially items 6 through 11.

The quantitative information available for the systems described in this section, basically direct or elapsed time savings, is only part of the picture of their true cost-effectiveness. In each case, a number of intangible benefits, including competitive advantages and the ability to improve products and reduce their costs, are highly significant but not quantified. In fact, in spite of the impressive time savings that have been observed with these systems, some authors believe that the development and use of existing computer graphics systems can only be justified by taking into account the intangible benefits; direct time benefits are not believed to be enough to justify the system costs that have been involved in existing systems (9, 14).

The fact that computer graphics development is continuing in industry indicates that it is the view of management that these systems are significantly cost-effective, although whether these gains are perceived as short term or long term is unclear.

A BENEFIT MODEL

A useful mechanism for summarizing the preceding sections is a simple benefit model. This model shows the basic relation that exists between interactive computer graphics costs and benefits:

$$TB + IB + DCB$$

where

TB = total benefit of using interactive computer graphics,
IB = other intangible and indirect cost benefits including benefits due to elapsed time savings, to better and/or more economical products or alternatives, and to efficiencies in the design process, and
DCB = direct cost benefits.

IB is, by its very nature, difficult to quantify and clearly will vary from application to application. It is obvious, though, that IB can be very large in cases where the product has a high value, leading to meaningful savings when the product is improved. If one considers the products to which production interactive computer graphics systems have been applied (automobiles, airplanes, fractionating towers, airport design), it seems clear that the value of the product is implicit in the decision to use the technique.

DCB, in dollars per year, is somewhat simpler to quantify.

$$DCB = T\left[W_m - F\left(W_g + VCC + \frac{ACC + FC}{D}\right)\right]$$

where

- T = total time spent on application without graphics, hours/year;
- W_m = wage rate for manual work in application area, dollars/hour;
- W_g = wage rate for work at graphics console in application area, dollars/hour;
- F = time saving rate using graphics = (time using graphics)/(time using manual methods);
- VCC = hourly variable computer-related and console-related computer costs for graphics applications, dollars/hour;
- ACC = annual fixed console-related and computer-related computer costs due only to graphics, dollars/year;
- FC = annual fixed costs for graphics software and overhead, dollars/year; and
- D = maximum console usage rate, hours/year.

D applies to all applications using a system, whereas the formula itself refers to a single application. Clearly, $D \geq TF$.

Although we stress that it is erroneous to consider only DCB in evaluating a graphics system, the expression for DCB is useful in that it illustrates the following points:

Due to the fixed costs, which include developments and can be very high [O'Neill (14) estimates 60 man-years of development effort in the McDonnell Douglas system], the total time spent on work in graphics application areas (T) must be high to ensure positive cost benefits.

The time saving rate using graphics must be significantly less than (manual wage rate)/(graphics wage rate + fixed and variable console costs) or

$$F = \frac{W_m}{[W_g + VCC + (ACC = FC)/D]}$$

If the rate is equal to or greater than this quantity, cost benefits will be negative. As manpower costs increase over time, the required value of F will increase, and more systems will become cost-effective.

Console costs can be critical. Assuming that console and computer costs continue to decrease, as they have in the last 3 or 4 years, more and more graphics applications will change from negative to positive direct cost benefits.

As more experience is gained with the design and implementation of graphics systems, the fixed costs (FC) of these systems can be expected to decrease. The availability and use of standard packages of system and utility programs will also cause a decrease in fixed costs. These changes will have a direct positive effect on the direct cost benefits of graphics systems.

In summary, based on this simple model of the direct cost benefits of computer graphics systems, graphics systems can be expected to become increasingly cost-effective as (a) manpower costs rise, (b) console costs decrease, and (c) development costs decrease.

EXISTING SYSTEMS WITH APPLICATION TO TRANSPORTATION AND PLANNING

The previous sections have reviewed the general state of the art in interactive computer graphics and have, it is hoped, given the reader an understanding of the conditions under which the concept is a cost-effective one. Before we go on to draw any conclusion on the relevance of this technique to the field of transportation, however, it is first useful and necessary to review the state of the art of computer graphics applications related to transportation and planning. These systems either are at the proposal state, have only been used in a test or prototype environment, or are in production use although no information on their cost-effectiveness is available.

Highway Design Systems

The first highway plotting using digital plotters in the United States was done at the Department of Civil Engineering at M.I.T. in 1961 (20). Since that time, a large number of programs have been developed, and many organizations have incorporated comprehensive plotting packages into production usage. Packages normally include geometry, profile, and cross section routines, and, in addition, many include network and traffic plotting capabilities. Existing packages are not interactive, and usually the output device is a paper plotter. Organizations using these capabilities include nearly every state highway department and many consultant firms who do highway design work.

The earliest highway perspective plotting was performed by Nordick, a European engineering firm, in the early 1960s. In recent years, a number of organizations have developed perspective drawing programs, some of which include a "scene walking" capability that permits a user to "drive" along the proposed roadway in a simulated fashion.

At the present time, interactive graphics systems for highway design are in the proposal, system design, and prototype stage. The proposal for the California Division of Highways systems (2) contains the following summary of estimated annual savings for a system of 166 storage device CRT terminals:

Item	Time (million hours)
Present design hours	4.15
Design hours with graphics	3.44
Design hours saved	0.71

Item	Cost (million dollars)
Wages saved ($7.25/hour)	5.15
Additional computer charges	1.95
Annual cost savings	3.20
Hardware acquisition	1.48
System development	0.58

In addition to the time savings quantified, additional savings due to higher quality design were predicted but not quantified.

Urban Planning

The most significant area in which computer graphics has been applied to urban planning problems is computer mapping. A survey of the systems available is given in Goldstein, Wertz, and Sweet (7). All systems surveyed were non-interactive. A number of research-oriented or prototypical interactive graphics systems with applications to urban planning now exist. Some of the more interesting of these follow.

DISCOURSE—This is a system (15) that allows the planner to describe an area divided into a grid by specifying the attributes of the cells. Once the area has been described, the planner can select subsets of cells that meet any number of conditions, such as having the value of specified attributes in given ranges and being adjacent to a particular kind of cell. New attributes can then be assigned to these subsets. Because attributes can represent such things as single-family housing construction and transit stations, the planner can propose changes to his analysis area, investigate the effects of these changes, and then accept or reject them. The planner is able to do this in an interactive computer-aided mode. In the first version of DISCOURSE, graphics provides just one of the aids available in the system: a "map" of the analysis area showing the values of a single attribute for each cell. Compared with normal planning practice, graphics is definitely downgraded in the system. The developer of DISCOURSE feels that this is justified, saying that "its graphics are intended only for the purpose of presentation—not analysis."

As a reaction to typical urban planning, which is often highly graphics-oriented, this approach to a system designed to improve the planning process appears to be warranted.

URBAN5—The purpose of this system (13) is to provide intimate communication between an urban designer and a machine, so that an evolutionary process can occur and so that the machine eventually will exhibit a kind of design intelligence, reflecting the methodology of a specific user. URBAN5 is therefore designed to study the artificial intelligence possibilities of a computer applied to urban problems. This highly experimental system allows the user to design within a three-dimensional rectilinear space and keep track of its own and user-supplied criteria such as maximum number of vertical surfaces in shadow and incompatibility of education and industrial spaces in the same location. Although the system was successful in providing a highly interactive system for the design of spaces, it was found to lack the generality necessary to be a true learning system.

URBAN COGO—This system (21) provides an urban information system based on such geometric objects as parcels, blocks, regions, and networks. Each of these objects, and a number of simpler ones, can be described with a user-generated set of attributes, such as number of buildings and number of families on a parcel. The graphical capabilities of the system include the following:

1. Graphical output capabilities, including both soft- and hard-copy displaying of objects or groups of objects with or without translation, rotation, or magnification, density mapping, selective mapping, and detailed mapping with full annotations; and
2. Graphical input capabilities by digitizing on a display screen or digitizing from hard copy on a flat-bed plotter digitizer.

URBAN COGO is designed to provide the base and direction for urban information systems of the future.

Santa Clara County Planning Department—In cooperation with IBM and the city of San Jose, the Santa Clara County Planning Department is developing a system of interactive computer programs for the prediction of the spatial distribution of households and commercial establishments (4). The system includes econometric, demographic, and location models that operate on a common data base. The graphical capabilities allow the user to display portions of the data base in a number of ways, including numerical listings on a CRT and analysis area maps with user-specified variables or operations on variables displayed for each analysis zone. When data are displayed numerically, they can be modified using the light pen and CRT keyboard. Development is continuing on additional submodels and expanded display capabilities.

The system is being used to study and evaluate the urban development policies of the local governments in Santa Clara County. This use is providing significant insight into the requirements for an interactive model system as a tool for regional planners.

Transportation Planning Studies

A pioneering use of CRT graphics was the "cartographatron" developed for the Chicago Area Transportation Study to display such transportation data as trip and locations and desire lines of area trips superimposed on an outline map of the Chicago metropolitan area. This device was operational as early as 1959 (3). In spite of this early beginning, computer graphics has played a relatively small part in transportation planning studies. Some use of the printer plotter mapping capabilities available in SYMAP exists (5). The paper plotter network displays available in the Bureau of Public Roads urban planning package of computer programs are used by a number of studies (26). Both of these applications are non-interactive. No cases of the use of interactive graphics by transportation studies are known.

Transportation Analysis Research

Although interactive graphics is not now being used in a production environment by transportation planners, research in this area has been very active in recent years.

In particular, the work done at the University of Washington, M.I.T., and the University of Illinois at Chicago Circle is of interest.

At the University of Washington, the Urban Systems Laboratory has been active in using an ARDS CRT console and a time-shared computer to develop three sets of capabilities: (a) a network manipulator that allows networks to be built and modified (18), (b) a network generator that seeks to find the "best" set of network changes (19), and (c) a prototype of an interactive transit system analysis package (12). Of the three, the latter appears to be the one most advanced from the prototypical stage to the production-oriented stage.

The M.I.T. Department of Civil Engineering has been experimenting with interactive graphics, as well as with ways to use a small computer with graphics capabilities as an engineering input-output station for the last 6 years. Before transportation application work was done, Foster (6) developed communications capabilities that allow the transfer of data between an IBM 1130 and an IBM 360. Stotland used these capabilities to send graphics information specified in 360 programs to an 1130 for plotting. At the 1130 end, the user can specify the plotting device but has little additional control of the picture produced (24). Silverstone and Mumford used Stotland's routines to produce network and desire-line displays generated in ICES TRANSET, a traffic assignment subsystem (23). Also, they experimented with the dynamic display of transportation data. For example, the speed on links of a network was displayed using dashed lines that "moved" from origin node to destination node at varying speeds.

Pradas-Aracil and Blumsack have developed a general system for the display of points and lines in one, two, or three dimensions (1, 16). This system allows files of n-dimensional points, and lines connecting them, to be read, edited, transformed, ordered, and stored for use in graphical outputs. Graphics can be drawn with a number of options, including specifying the dimensions of the axes, fitting of regression lines to points, and showing envelopes of maximum and minimum values. Although general in terms of the types of point and line graphs that can be obtained, this system does not include special features for such typical transportation graphics as network flow maps. The output device for which the system is designed is an ARDS CRT terminal with a keyboard for user input of information. Experimentation with this system has indicated that simple, only partially interactive systems can have sizable direct cost benefits.

At the University of Illinois at Chicago Circle (10), the Department of Systems Engineering has begun the development of a general computer system for the interactive analysis of planning and transportation problems. The development philosophy is similar to M.I.T.'s in that the emphasis is on a number of analysis models that operate on a common data base, as well as a package of graphical capabilities that can be used to display information from the data base. The initial version of the system has been named INTRANS (Interactive Transportation Analysis System). It currently includes, as one of a number of potential models, a subsystem for data analysis named BROWSE. This subsystem allows spatial variables to be displayed on a map of the study area and frequency and functional variables to be displayed in mathematical plots. Data base variables may also be transformed using mathematical operations. Future plans call for the addition of new subsystems to INTRANS, to include transportation analysis capabilities, to operate both on real data, such as zonal populations and commercial development, and on network data, such as multimodal transportation networks.

Summary

The use of graphics in the field of transportation can be summarized as follows:

1. There have been a variety of applications developed, production systems tend to be static rather than interactive, and those interactive systems that do exist have tended to be experimental in nature;

2. A great many of the applications are geographically based;

3. Experimentation on cost-effectiveness of the use of graphics in the transportation field has been encouraging (work at M.I.T. has indicated that even rather simple static systems can have cost benefits to the user); and

4. Some very useful first steps have been made, but a comprehensive interactive graphics system for transportation is still a long way off.

CONCLUSIONS

The underlying notion of our research has been to consider the relevance of interactive computer graphics in transportation analysis. The method chosen was a comprehensive look at trends in the field in general and existing operational systems both within and without the transportation field. Based on this work the following conclusions were reached.

From a direct cost viewpoint, interactive computer graphics has historically been uncompetitive, even though existing production systems demonstrate that man-time savings are easily achievable. This stemmed from the very high hardware and software costs present in the field. However, hardware costs are falling as are software costs (in fact, software now often exists for particular applications). At the same time man-time costs are rising. Therefore, it is expected that computer graphics will become more competitive in the future.

The intangible benefits of getting the job done better and more quickly must be considered in evaluating an interactive computer graphics system. This has historically been the case, in that existing production interactive graphics systems have been developed where the "product" was one of high value, due either to high costs (as in the aerospace industry) or to high volume (as in the automotive and oil industries). Because the product in transportation systems analysis is typically extremely high-valued and also very long-lived, the motivation for using techniques that will enhance the product clearly exists. An added incentive to the development of transportation computer graphics systems is the high social costs associated with "second-best" transportation systems, which are so closely meshed in our society.

A survey of proposed or prototype transportation graphics capabilities reveals that a substantial amount of work has been done in many aspects of the field. The potential for effective use of interactive computer graphics in transportation has been demonstrated. In addition, it does not appear that graphics capabilities need be fully interactive to be useful and cost-effective in all applications in the transportation field. Experimentation with the prototype M.I.T. transportation graphics indicates that simple, only partially interactive systems can have sizable direct cost benefits.

As indicated by the formula for direct cost savings, the amount of analysis time that lends itself to graphics applications is a critical variable in the determination of direct cost benefits. Existing industrial graphics systems have all been developed by very large firms with very large engineering staffs. It therefore becomes necessary at some point in time to approximate the potential use of a large-scale transportation graphics system. Initial thinking indicates that the system may require use by a substantial number of transportation analysts in order to have positive direct cost benefits. This argues for a coordinated approach to the development and use of an interactive graphics system by the transportation planning community.

In summary, the direct cost picture in interactive graphics is changing for the better, as hardware and software costs fall and man-time costs rise. The public climate is such that arguments for mechanisms for developing better transportation alternatives are likely to be heard. Substantial progress has been made by individual researchers in the field who have demonstrated that interactive graphics can be effective in transportation systems analysis. We feel that the time is ripe for the profession to take a coordinated cooperative look at the use of interactive graphics in transportation systems analysis, a look that will hopefully lead to useful modular interactive packages for the field at large.

ACKNOWLEDGMENTS

This paper was developed as part of a research project on computer graphics for transportation analysis carried out in the M.I.T. Department of Civil Engineering under the sponsorship of the Transportation Systems Center, U.S. Department of Transportation. The authors wish to acknowledge the helpfulness of Harold Miller, Peter Segota, and Richard Cox of TSC who provided technical monitorship of the project. The authors would also like to acknowledge the efforts of Daniel Roos who helped by reviewing and suggesting revisions for this paper.

REFERENCES

1. Blumsack, S. L. 1130 Graphical Evaluation System. Dept. of Civil Engineering, M.I.T., Cambridge, MS thesis, 1971.
2. Interactive Graphics for Highway Design. California Division of Highways, 1971.
3. Chicago Area Transportation Study Final Report, Volume I: Survey Findings. C.A.T.S., 1959.
4. Cristiani, E. J., et al. An Interactive Facility for Examining Development Policies of Local Government. Presented to Ninth Annual Allerton Conference on Circuit and System Theory, Urbana, Ill., 1971.
5. Fisher, H. T., et al. Reference Manual for Synographic Computer Mapping (SYMAP), Version V. Laboratory for Computer Graphics, Harvard Univ., Cambridge, Mass., 1968.
6. Foster, A. T. The Use of Small Computers as Engineering Input-Output Stations. Dept. of Civil Engineering, M.I.T., Cambridge, PhD thesis, 1969.
7. Goldstein, H., Wertz, J., and Sweet, D. Computer Mapping: A Tool for Urban Planners. Urban Studies Center, Battelle Memorial Institute, 1969.
8. Green, R. E. Computer Graphics in the United States. Computer Aided Design, Vol. 1, No. 3, 1969, pp. 33-38.
9. Green, R. E. Computer Graphics—A Status Report From the U.S.A. Computer Aided Design, Vol. 2, No. 4, 1970, pp. 3-9.
10. Gur, Y. A Computer System for Interactive Analysis and Planning of Urban and Transportation Problems: Progress Report No. 1. Dept. of Systems Engineering, Univ. of Illinois at Chicago Circle, 1972.
11. Jacks, E. L. A Laboratory for the Study of Graphical Man-Machine Communication. Proc., Fall Joint Computer Conference, Spartan Books, Inc., Baltimore, 1964.
12. Koike, H. Planning Urban Transportation Systems: A Model for Generating Socially Desirable Transportation Network Configurations. Depts. of Urban Planning and Civil Engineering, Univ. of Washington, Res. Rept. 2, 1970.
13. Negroponte, N. URBAN5: An Experimental Urban Design Partner. In Computer Graphics in Architecture and Design, Yale School of Art and Architecture, 1969.
14. O'Neill, J. B. Computer Graphics as an Operational Tool. Presented to Computer Graphics 70 International Symposium, Brunel University, Uxbridge, England, 1970.
15. Porter, W., Lloyd, K., and Fleisher, A. DISCOURSE: Computer Assisted City Design. In Computer Graphics in Architecture and Design, Yale School of Art and Architecture, 1969.
16. Pradas-Aracil, F. A. A Graphical Evaluation System. Dept. of Civil Engineering, M.I.T., Cambridge, MS thesis, 1970.
17. Prince, M. D. Interactive Graphics for Computer-Aided Design. Addison-Wesley, 1971.
18. Rapp, M. H. Interactive Editing of Transportation Networks. Depts. of Urban Planning and Civil Engineering, Univ. of Washington, Res. Rept. 4, 1970.
19. Rapp, M. H. The Planning of Node-Oriented Transit Systems. Depts. of Urban Planning and Civil Engineering, Univ. of Washington, Res. Rept. 3, 1970.
20. Roberts, P. O., and Suhrbier, J. Highway Location Analysis: An Example Problem. M.I.T. Press, Cambridge, 1966.
21. Schumacher, B. URBAN COGO Users Guide. Urban Systems Laboratory, M.I.T., Cambridge, First Ed., 1970.
22. Siders, R. A., et al. Computer Graphics, A Revolution in Design. American Management Assn., New York, 1966.
23. Silverstone, S. M., and Mumford, J. R. Development of Graphic Communications in a Computer-Based Transportation Planning Process. Dept. of Civil Engineering, M.I.T., Res. Rept. R68-50, 1968.

24. Stotland, H. C. Remote Graphical Capabilities for an Engineering Input-Output Station. Dept. of Civil Engineering, M.I.T., Cambridge, MS thesis, 1968.
25. Sutherland, I. E. SKETCHPAD—A Man-Machine Graphical Communication System. Proc., Spring Joint Computer Conference, Spartan Books, Inc., Baltimore, 1963.
26. Urban Transportation Planning: General Information and Introduction to System 360. Federal Highway Administration, U.S. Dept. of Transportation, 1970.

INTRANS AND BROWSE: AN INTERACTIVE GRAPHICS SYSTEM FOR PLANNING RELATED DATA ANALYSIS

Yehuda Gur, University of Illinois at Chicago Circle

This paper describes the Interactive Transportation Analysis System (INTRANS). INTRANS is a man-computer interactive graphics system designed to serve transportation and urban planners. Currently, INTRANS is being used mainly for analysis of spatial data. Its major use in the future will be as a skeleton for interactive planning models. INTRANS operates in the Computer Center of the University of Illinois at Chicago. It uses an IBM 370 as the central processor. The paper defines the need for interactive graphics in planning, the areas where its use is likely to be effective, and the major problems in its implementation and use. It then describes the design of INTRANS, its major elements, and how the system as a whole is being used. The paper concludes with a description of plans for development for both the short term and the long term.

•DURING the last 2 decades, a new type of sociotechnological planning process has emerged. It was initiated in the mid-1950s with the urban transportation planning process (1); during the 1960s, it was spread to various urban growth and activity allocation models [e.g., Boyce and Day (2)]. Currently, this type of planning process is applied to other urban subsystems as well as in other related areas.

This process is characterized by the set of tools that are inherent in its application. These tools are called here, in aggregate, the data-models-computer (DMC) system. Large volumes of data are collected and produced by the planning process. These data are used as a basis for a detailed quantitative description of the region. The data are analyzed by a set of mathematical models including prediction, cost, and network analysis models.

The handling of data as well as the calibration and application of the models is possible only through the use of computers. As a matter of fact, specific analysis procedures have always been designed by considering the limitations of available computers and have been improved with advances in computer technology.

The DMC system has been continually growing and improving with increased volumes of data and better models and computers. Its potential to supply knowledge on our environment and to support planning activities is quite large. However, the utility of the DMC system is curtailed by the difficulties in communications between it and the human users. Improving communication and interaction between people and the DMC system is a most cost-effective step in the current state of affairs.

CHARACTERISTICS OF INTERACTIVE GRAPHICS SYSTEMS

A promising way to improve man-DMC communications is by the use of man-computer interactive graphics systems (IGS). In these systems, the user and the computer communicate in real time; a large part of the information transfer is done graphically through a cathode ray tube. In this paper, the development and performance of such a system are described.

A number of advantages of IGS are obvious. Most important is the significant reduction in time consumption. Real-time interaction cuts the overhead time spent in punching cards, submitting a job, and waiting for the output to return. Graphic displays enable perception and comprehension of information much faster than do other methods; this is particularly true in spatial analysis where both the intensity relationships and the locational relationships between variables are important. Direct access to the computer enables one to assign to it a large number of relatively small tasks that otherwise would be done manually. It also enables immediate access to large amounts of data and various model programs that are stored in the computer and its peripheral equipment.

Potential Uses of IGS in Planning

Access to Information—Because of its characteristics, IGS might be used effectively in many phases of the planning process. Easy access to large amounts of information may enable planners and model builders to examine raw data and final, as well as intermediate, outputs of various models much more thoroughly than currently feasible. This may result in a significant reduction of errors, better understanding of the models, model improvements, more realistic assessment of their limitations, and reliability of model outputs.

No less important is the potential use of IGS in presentations to the general public and to decision-makers. Easy access to large amounts of information (and, possibly, models) will enable proper immediate response to a wide range of inquiries and the development of meaningful and factual discussions in many, possibly unexpected, directions. This will be a significant improvement on the existing situation where presentations are limited to material that is fully prepared beforehand.

IGS can be used effectively to access a "bank" of base information on a region. Such information could be stored permanently on the IGS files and then retrieved in response to users' requests. Flexibility in choosing the format and content of the retrieved information and hard copy capabilities, which are usually available, are very useful for this purpose. Using IGS in this framework can substantially increase the quality and speed with which planning agencies can respond to information requests, both internal and external.

Editing—Another potential application of IGS is in data editing, particularly in cases where human judgment is required in the editing and where graphical display of the stored information might be of help. Such a case is, for example, the editing of coded networks (5). Here, the required record keeping, cross references, and routine calculations can be done by the computer, relieving the analyst of these error-prone activities. This is additional to the advantage of being able to observe graphically the parts of the network that are of interest, in the manner in which they are coded.

Interactive Planning and Design—A family of applications in which IGS seems to have a large potential is in plan formulation, plan modification, and design. Here, the sophistication of the systems can vary significantly. In one extreme, IGS can be used as an editor, enabling easy storage, retrieval, and changes in plan details and not much more. Systems could be somewhat more sophisticated, where, through the use of various cost and prediction models, they would supply the analyst with estimates of required resources and impacts of alternative plans. They may also make routine calculations and fill standard details. At the other extreme, interactive graphics planning systems may be quite sophisticated. By using programming techniques, they may find optimal plans, in response to the analyst's specification of the solution space, the constraints, and the objective function. In this field of applications, IGS seems to be most useful in sketch planning where rough examination of many alternatives is required and in introducing and analyzing small-scale local improvements in an existing plan where the marginal impacts can be estimated by simple models without reanalyzing the whole plan.

Difficulties in Use of IGS

The list of potential uses can be made much longer. However, there are a number of constraints on the effective use of IGS that strongly limits its applicability. Some of the constraints are inherent in the characteristics of IGS and human users; others are temporary and likely to be relieved in the future. The constraints are due to the characteristics of the human analyst, available computer technology, the type of problems in transportation planning, and the institutional structure of planning activities.

First, the large power of IGS and the interactive environment impose a high load on the analyst using them. The rate of information being transferred to the analyst and the rate of decisions that he has to make increase significantly. Effective use of IGS will strongly depend on how well it will be adapted to human constraints. For example, to make the use of IGS easy, the user would tend to make his commands simple. At the same time, in order to maximize user control of the IGS operation, a large set of commands, hence a complicated command structure, is necessary. This trade-off between simplicity and flexibility is only one of the difficult but necessary decisions that have to be made in order to make IGS best fit human characteristics. It is likely that new planning strategies, new techniques, and special training will be required before IGS can be implemented effectively. Only actual experience can point out optimal IGS designs.

Constraints due to computer technology are numerous. First, IGS demands, by definition, a short response time (maximum of a minute or so). If we consider the speed of available computers, this limitation excludes a large number of problems and existing models. It is likely that many models will have to be reformulated (e.g., by using marginal analyses) in order to be used by IGS. The need for fast "number crunching" and access to large data files implies the need for large computers. At the same time, special technical needs of IGS cause difficulties in running these jobs in a multiple-job environment typical to large computers. Some of these problems can be solved by dedicating full installations (on a full- or part-time basis) to planning IGS. (This has been done by Design IGS, being used in the automobile and aerospace industries.) However, transportation planning is currently being done by many separate, small- or medium-sized agencies using different computing facilities.

Good cost-effective solutions to these problems are yet to be found. However, continual improvements in computer technology allow many of the problems to diminish quickly with time.

The nature of activities in transportation planning imposes some limitations on IGS applicability. Many of the activities in the planning process are not repetitive or are repeated very infrequently. This implies that activity-specific computer programs will not be used very frequently. Because the overhead in designing and implementing any IGS is quite high, it might not be cost effective. No less difficult a problem is the fact that many of the activities in transportation planning have not been standardized yet. They differ significantly from place to place and from problem to problem and are being modified continuously. This implies that it will be difficult, if not impossible, to implement a general transportation planning IGS. In a number of problems where there is standardization, e.g., geometric design, these problems are less serious. [Such a system was designed by Beilfuss, Dwyer, and Phillips (4).] With increased standardization in the field, e.g., the BPR system (6) and the HUD transit package now under extensive development, these problems might be diminished but are unlikely to vanish.

In the following paragraphs, the development and performance of an IGS, called the Interactive Transportation Analysis System (INTRANS), is described. Many of the considerations described affected the design and implementation of INTRANS. In the description, points where trade-offs had to be made are specified, and the reasons for the specific decisions are presented.

DESIGN AND IMPLEMENTATION OF INTRANS

The Computer System

At the very start of the project, it has become clear that the scope of INTRANS will depend strongly on the available computer system. More than that, the resources available to this project preclude any substantial investments in computer hardware. Thus, the problem for this project has been to find the most powerful yet accessible computer and design the IGS around it.

The system that has been chosen is an IGS developed by the computer center at UICC. The system is shown in Figure 1. This system, together with the available software, has the following characteristics.

1. Calculations, data storage, and data management are done by the IBM/370 under OS. This enables access to most of the equipment around it, in particular, disk packs for data and program storage. It also enables the use of the full capacity, power, and speed of the 370 CPU and core memory.

2. The computer system provides for the use of many higher level programming languages, as well as many of the available library programs.

3. The IBM 1800, with the Channel, is used as a controller for fast data transfer. Its potential use as an auxiliary processor has been recognized but not yet implemented. Alternatively, its existing functions can be performed by less "intelligent" and much cheaper devices. (It is expected that by mid-1973 INTRANS will operate through standard high-speed time-sharing communication devices, with no need for the 1800 computer.)

4. The Tektronix storage display tube has a screen measuring 16 x 22 cm. It does not enable dynamic images to be displayed (as is possible with some other display tubes). This limitation pays off in a much lower hardware cost, lower load on the computer during display, and higher resolution.

5. The basic software for communication between the Tektronix and the 370 includes Plotter-type commands for creating graphic displays and routines for two-way transfer of character strings. All these routines can be called by FORTRAN.

6. The system operates under TSO (time sharing). This ensures that, while the user "scratches his head," the load on the 370 is minimal. The option to operate in batch mode (where the program resides permanently in core) is available. It might be useful for smaller computers or in especially large problems. This attribute substantially decreases the cost and operating difficulties of INTRANS as compared to similar IGSs.

7. Currently, the user communicates with the program through a keyboard. A joy stick with a cursor is being added to the system for efficient graphic input.

System Design

Given the capabilities of the computer system and the need for and problems in using IGS in transportation planning, the desired characteristics of INTRANS can be specified.

Real-World Applicability—The capabilities of the available computer system seem to be large enough to enable the use of INTRANS in relatively large problems. Thus, the development of features that might have immediate applications in full-scale, real-world problems has been stressed.

A System—Not a Model—INTRANS has been designed as a system that can support a large set of interactive graphics models rather than as one specific model. This has been done to enable easy adaptability to the wide range of planning applications, to study man-IGS interaction characteristics in simple applications, and to enable continual growth with increased experience. INTRANS consists of the elements that must be included in any planning-oriented IGS. Each separate model is written using the elements of INTRANS, thus ensuring relatively easy future aggregation with other models.

INTRANS is designed to include the elements that are most difficult to program. Thus, adaptation of specific models can be done by people without advanced knowledge in computer programming.

BROWSE—A Model for Data Analysis—A natural first step in developing a planning IGS seemed to be a model for interactive analysis of existing data. A model called BROWSE has been developed for this purpose. It consists of the basic elements of INTRANS with minimal additions. Besides being used for the many needs of data analysis, BROWSE may be used to study characteristics of man-IGS interaction and to evaluate the outputs of other interactive models that will be implemented in the future.

Elements and Structure of INTRANS

When we consider the general type of problems that INTRANS is likely to be used for, the programming problems, and the requirements specified, it becomes apparent that INTRANS has been designed in detail. The major elements in the system are shown in the lower part of Figure 2. In the design of each of these elements, there is a compromise among the consideration of simplicity, minimal core, speed, flexibility, and generality. The following sections discuss the elements in detail. Only experience will show whether the options chosen are the right ones.

Geographic Identification Methods

A major decision in performing spatial analyses is how locations are to be defined and identified. This decision has implications for the structure of data files as well as for details in the logic of many computer programs. INTRANS is an analysis system intended to be used in many different areas and for many different problems. Thus, it should use a specific identification method rather than a unique system to which data management and program logic will be adapted.

Geographical identification is needed for points (e.g., location of a road intersection or a school) and for areas. For points, it is required that any point in the area be identified uniquely. The most widely used method is the Cartesian coordinate system. More than that, this is the method used in the existing interactive computer software. Thus, it has been adopted by INTRANS as the only reasonable choice.

For areas, it is required that the whole region be subdivided into mutually exclusive and conclusive analysis subregions, called zones, and that each zone be identified uniquely. The two major alternative methods are a general system and a grid. In the general system, boundaries of zones, their size, and their shape are specified at will, usually by considering the needs of the analysis; zones are identified nominally. In the grid method, grid squares are considered to be zones. By choosing an origin, direction, and scale for the grid, the exact boundaries of each zone are fixed. Zone identification number can be easily related to its location.

After these alternatives are compared, the choice has been to use a grid system. Following are the major factors that affected the choice.

Simplicity and Economy—In using a grid, much information is available on a zone if its number is known. This includes location of centroid, boundaries, adjacent zones, and area. The coding, storage, and analysis of this information in the general system are expensive in manpower, core, and computer time.

Display Clarity—When a grid is used, the density of zone centroids is uniform. This enables the creation of effective displays of areal distributions by relating the amount of light in each grid to the intensity of the corresponding variable, for example, by using different characters on a "gray scale." (See, for example, SYMAP.) When the general system is used, the distribution of light intensity depends largely on the areas of zones (or distance between centroids).

Data Availability and Use by Planning Agencies—In general, the many agencies that collect and analyze data use different and noncompatible areal units, e.g., census tracts, zip codes, communities, towns and cities, and school districts. From this point of view, there is an advantage to the general system. However, when this factor is considered, it should be noted that many planning agencies, including those in the Chicago region, use a grid as a basis for zone definition.

Efficiency in Analysis—The major drawback in use of a grid system is its inefficiency in analysis. In the general system, it is possible to relate zone sizes to analysis needs, defining small zones only where details are needed. In this way, it is possible to get the required level of detail without a substantial increase in the amount of calculations. This freedom is not available in the grid system.

After all these factors were weighed, it has been judged that the advantages of the grid system outweigh the disadvantages. It is possible to use INTRANS for many problems where a general areal ID system is used. However, in general, this might require special adaptation, and it does not give all the capabilities available to the grid user.

DATA MANAGEMENT SYSTEMS

The important function of the data management system (DMS) in INTRANS cannot be overstressed (3). First, compatibility between various models depends largely on compatibility in data handling. Second, INTRANS is intended to handle large amounts of data. The efficiency and speed of the system depend largely on the efficiency of the I/O operations. A third reason for the importance of the data system in INTRANS is the need for advanced programming techniques. FORTRAN is quite inefficient in terms of core in its nonstandard I/O codes when compared to what can be done in machine-oriented languages. Thus, it is important to relieve the application programmer from the need to program this element.

Functional Organization of the Data

The Study Area—This is the major element in the structure of DMS. Most of the likely analyses by INTRANS will be done on one study area at a time. A study area is represented by a grid of given dimensions (NCOL by NROW) and a given scale (size of a grid). Each INTRANS data set refers to one such study area.

Files—Each study area includes one or more files. Each file stores a cross-sectional picture of the area. For example, different files might relate to different years; alternatively, files might relate to different alternative plans. It is possible to specify summary files that include variables of each of the separate files.

Variables—Each file consists of a number of variables. Variables in different files might have identical names. This is quite efficient for making the individual programs file-independent. Each variable consists of a number of elements; it is possible that different variables within a file have different numbers of elements. A variable might be, for example, population. Each element gives the population in one of the zones. Similar to most variables currently used, this variable has NZONE elements. However, another variable, say trip length distribution, might have a different number of elements. I/O operations are done on the level of variables; i.e., all the elements of a variable are stored and retrieved simultaneously.

Technical Details

DMS uses the structure of partitioned data sets. Each study area is one data set; each variable is a member. Each variable, as well as the data sets, can be labeled with up to 56 characters. Labels are stored in the data set directory. Variables are stored in binary form; they are retrieved directly into the required location without intermediate buffering. There are practically no limitations to the number of variables that can be stored.

DMS programs are used off-line for creation and maintenance of data sets. During interaction, the user can access these programs in order to create or delete variables. The display and compute routines, as well as other models, use DMS programs to retrieve and store variables.

DISPLAY ROUTINES

The quality of the interactive system being constructed depends, to a large extent, on the quality of displays it can produce. The clearer the displays are, the easier it will be for the user to perceive the information they contain. This will determine both the efficiency and the utility of the system.

Good displays are especially important if the system is to be used to transfer information to nontechnicians. In this case, the system will not be usable unless the displays are easily understood. They should be understood even by people who do not have advanced training or experience with the system. (It is not intended that nontechnicians will actually operate INTRANS with no training. However, they should be able to get clear answers to inquiries given to an operator.)

Thus, INTRANS includes the capability to create many types of displays and to closely control the design of these displays.

Types of Displays

The following types of display are available in INTRANS (Appendix).

Map—Maps displaying the areal distribution of intensity of zonal variables (on a grid) can be produced. This type includes three options:

1. Triangles on a 10 x 15 grid with a triangle whose side is proportional to the intensity of the displayed variable in the center of each grid,
2. Numbers on a 10 x 15 grid with one or two 3-digit numbers showing the intensity of the variable(s) within each grid, and

3. Symbols on a 40 × 60 grid with the intensity of variables in each grid described by a character as in SYMAP.

Symbols are used mostly for describing general trends in an area, whereas triangles and numbers "zoom in" for further details. The user can overlay the display with any of a number of precoded maps, showing locations or networks or both.

Distribution—Distributions display frequency or accumulative distribution of variables.

Function—Functions display functional relationship between two variables.

Distributions and functions can be described as a scattered diagram, polygon, step function, or column diagram. As an option, the user may get basic statistics (means, totals, correlations, regression lines) of the variables being displayed.

Controlling the Display Design

The display programs are called by the display command. Display command can be input either by the interactive user or by FORTRAN routines. By specifying the value of control variables, the user can specify the type of display, group, transform, scale, and bound of the displayed variables and thus control in detail the display design. Alternatively, the user can use very short commands, leaving the design of the display to the program.

THE INTERPRETER

The function of the interpreter is to accept user inputs from the keyboard (in the future, also from a joy stick) and change them into numbers that are used as values or addresses by the other programs.

The sophistication of an interpreter increases with the amount of flexibility given to the user. In a highly flexible system (e.g., interpretive languages), the interpreter should be very powerful, enabling the user to give almost any instruction he wants. In a more restrictive environment, the user has very few options, mostly requesting that certain routines be called. In such an environment, the interpreter is quite simple.

In INTRANS, the general tendency is toward a relatively restrictive environment. However, the interpreter that has been implemented is relatively flexible. This may even be some "overdesign."

The main reason for the designed flexibility is that INTRANS is a general system rather than a specific model. The interpreter is intended to serve almost any requirement by a specific model that is yet unspecified. It seems preferable to add some slack now, rather than be forced later to introduce additions that will probably cause incompatibilities.

Another reason for the relative flexiblity of the interpreter is that it was designed to answer the needs of the display routines in the BROWSE mode. It is not likely that many other models will give the user so much flexibility as is required for this specific application.

The interpreter is called by the main model program. It accepts a string of characters, which is called "command" from the user. The interpreter breaks a command into words and checks them for validity. Valid command words cause the interpreter to assign specific numbers to specific locations in an array. A nonvalid command word causes the command to be rejected. The application programmer is able to specify the command words, put constraint on their order, and specify the form in which the interpreted command will be transferred to his program's control.

The interpreter is capable of handling prespecified command words, free-format numbers, and names.

THE COMPUTATION COMMANDS

The computation commands are intended to give the interactive user the ability to formulate and perform computational procedures during interaction. It is expected that most of the procedures will be precoded as subroutines, and the user will have only to choose among them. The computation routines are intended to be used in the few cases where this extra capability is needed.

For example, consider analysis of data on zonal population and land use for a number of years. The analyst might be interested in seeing things like percentage of population increase, net and gross densities, intrazonal relations among various land uses, etc.

It will be practically impossible to precompute or code the routines to compute all the possible combinations. Using the computation command, the user can directly ask exactly what he is interested in.

A second example might be in evaluation. A user might be interested in formulating various measures of performance or weighing schemes. In many cases, it is much simpler to specify the measures during analysis rather than precode all the possible formulations. A typical computation command may look like

DDEN = POP1/RESA1 - POP2/RESA2

This calculates the net residential densities for two cases and the difference between them. The difference is stored in the variable DDEN for further reference. The computation commands use their own interpreter.

OTHER ELEMENTS IN INTRANS

Utility Commands

A number of utility commands are included in INTRANS and are directly available to any specific models. For example, SHOW DIRECTORY lists all the variable names in the data set, together with their labels; SHOW MENU displays all the valid command words. A data editor that enables the user to list and change variables and elements is also available.

Support Procedures

A number of batch processing routines are available for INTRANS users. First, there are programs for data maintenance and program library maintenance. Second, there are FORTRAN batch processing programs that simulate the interpreter and display routines. These programs are used in the debugging of new models. Third, there is a time-sharing version of INTRANS without graphic displays. It is accessible from remote terminals.

USING INTRANS FOR AN INTERACTIVE MODEL

INTRANS alone is a set of relatively independent routines. To activate INTRANS requires that all these routines be tied together with specific model routines into one system. Figure 2 shows the elements needed for making an interactive model that uses INTRANS.

The major added element is the model's control program. Usually quite short, the program's main function is to call the interpreter and the required routines and retrieve and store the data based on instructions received from the interpreter.

Each model has its own set of command words. They should be entered into the required interpreter arrays in advance, using a special program. Each model usually has its own computing routines.

CURRENT STATUS AND PLANS FOR FURTHER DEVELOPMENT

At the present, the programming of INTRANS and the model BROWSE for data analysis has been completed. Manuals for interactive users and application programmers have been written (8). An extensive data set on the Chicago area has been assembled and used for experiments in interactive analysis. Thorough testing of BROWSE has been completed. A number of relatively simple additional models are being developed. One is a model for study of the behavior of the gravity and opportunity distribution models. The second is a model for estimating the spatial characteristics of the effect of pollution emission from the area's expressway systems.

Figure 1. INTRANS computer hardware.

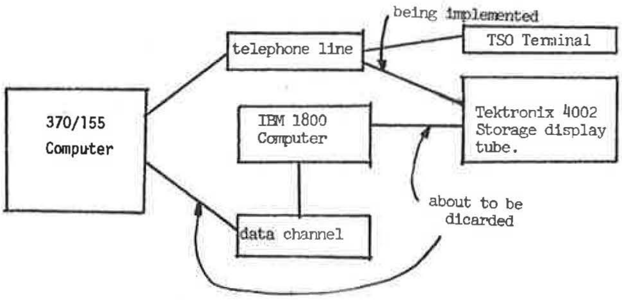


Figure 2. Structure of interactive model using INTRANS.

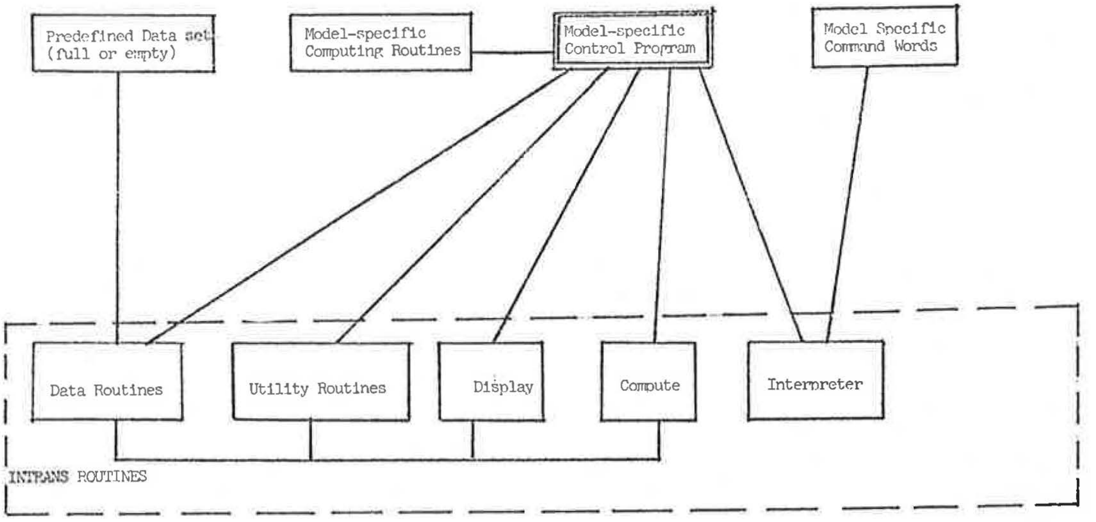
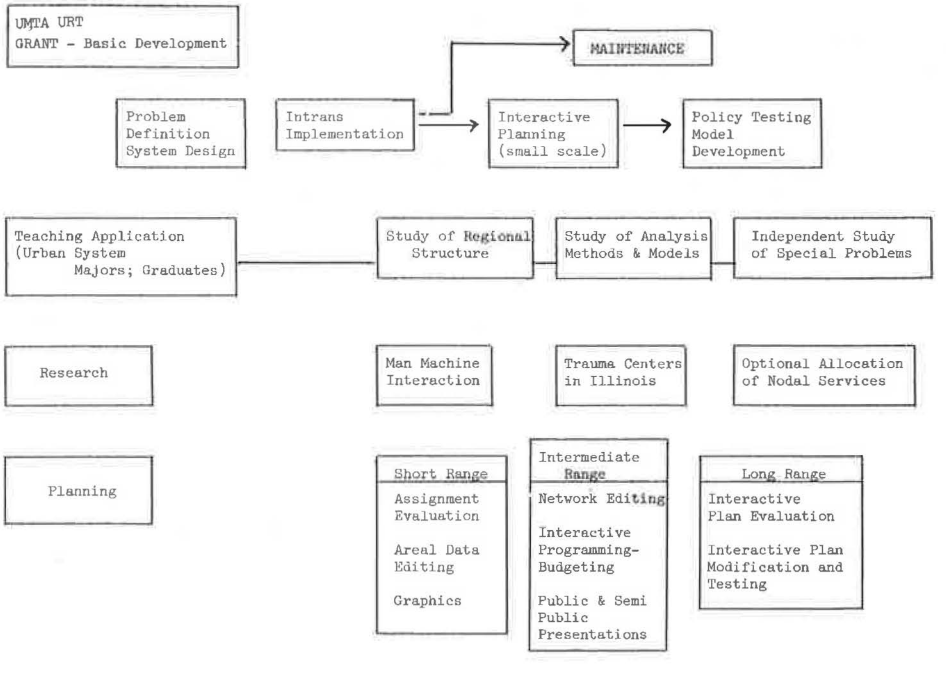


Figure 3. Tracks and major work items in INTRANS development.



Further development of INTRANS will be made in five parallel paths, as shown in Figure 3. This includes the development of INTRANS as a teaching tool, a research tool, and a planning tool. The two other paths are the maintenance and improvement of the existing system and the development of a general, large-scale policy-testing model (POTEM).

A Teaching Tool

At present, INTRANS may be used in teaching in a number of ways. First, students of urban systems may study structures of regions and relationships between intensities of various variables. Second, INTRANS may be useful in the study of graphic reporting techniques. Third, it is possible to use INTRANS as a laboratory for comparative analysis of the behavior of various models. With future availability of planning models, INTRANS may be used as a laboratory for interactive planning.

Limited experience during winter 1972 showed that effective use of INTRANS with the time constraints of a regularly scheduled course requires detailed development of laboratory manuals and problems. It seems that INTRANS can be used effectively mostly in advanced undergraduate and graduate courses because of the need for relatively wide background knowledge. Available time on the Tektronix also limits the amount of use of INTRANS for this purpose.

A Research Tool

INTRANS can provide substantial aid to research of problems with important spatial aspects. By using the available capabilities of the system, the user is relieved of much of the effort in data handling and analysis. This is additional to the ability to program specific problems for interactive data analysis.

Research projects that may use INTRANS are of a wide range of size and complexity. At one extreme, they may be student projects or independent study courses. At the other extreme, it may be possible to use INTRANS as an aid in large projects, particularly those that are data analysis oriented.

Developments in this area are expected to be initiated by researchers of specific problems rather than by the INTRANS development group. Thus, the extent and depth of use are unknown.

Currently, a number of projects are under way. First, there are three projects involving allocation of nodal services. One project is aimed at developing a multiple-root minimum tree-building program, which will be used in many optimal allocation problems. The second project deals with developing an interactive model for allocation of fire stations. The third project in this group uses INTRANS as an aid to study the allocation of trauma centers in Illinois.

A Planning Tool

Much effort has been spent to get planning agencies to use INTRANS as part of their activities. Only with experience in such applications will it be possible to develop the system to its full potential. In the immediate future, the most promising application seems to be browsing and editing of large data sets. Using the experience that will be gained in relatively simple applications will allow more sophisticated problems to be tried.

Currently, a study of the use of INTRANS by the Chicago Area Transportation Study is under way. It includes the experimental application of INTRANS for evaluation of a land use-activity data set and output of the transit assignment packages. Another project aimed to implement INTRANS as a "bank" for base information for the Illinois Department of Transportation.

ACKNOWLEDGMENTS

The research project described in this paper has been supported by a grant from the Urban Mass Transportation Administration, U.S. Department of Transportation, under the University Research and Training Program.

REFERENCES

1. Creighton, R. L. Urban Transportation Planning. Univ. of Illinois Press, 1970.
2. Boyce, D. E., and Day, N. D. Metropolitan Plan Evaluation Methodology. Institute for Environmental Studies, Univ. of Pennsylvania, 1969.
3. Oghigian, N. A Data Management System for Interactive Analysis of Transportation Problems. Univ. of Illinois at Chicago Circle, MS thesis, 1971.
4. Beilfuss, C. W., Dwyer, T. A., and Phillips, R. M. Interactive Graphics Roadway Design System. NCHRP Proj. 20-8, 1971.
5. Rapp, M. H. Interactive Editing of Transportation Networks, Report 4. Urban Transportation Program, Univ. of Washington, 1970.
6. An Introduction to System 360. Urban Transportation Planning, U.S. Department of Transportation, June 1970.
7. The DIME Geocoding System. U.S. Bureau of the Census, 1970.
8. Gur, Y. INTRANS-BROWZE User Manual. Dept. of Systems Engineering, Univ. of Illinois, Chicago, Tech. Rept. 2, 1972.
9. Haack, H. Geocoding at CATS—A Before and After Review. Presented at the 8th Annual Conference, URISA, Louisville, Ky., Sept. 1970.

APPENDIX

TYPICAL INTRANS DISPLAYS

The following figures are reproductions of photographs of the screen, taken with a 35-mm camera on black-and-white film.

Figure 4. Directory of file BDLAKECO. Socioeconomic data on Lake County, Illinois, 1 grid = 1 mile square.

```
SHOW DIRECT
BLAKECO
BLAKECO
BA•       ....  LAKE COUNTY ILL. SOCIO ECON. DATA
BAUTO6    ....AUTOS      AUTOMOBILES 1965
BAUTO9    ....AUTOS      AUTOMOBILES 1995
BEMFG6    ....JOB       MANUFACTURING EMPLOYMENT 1965
BEMFG9    ....JOB       MANUFACTURING EMPLOYMENT 1995
BLCOM6    ....ACRES      COMMERCIAL LAND 1965
BLCOM9    ....ACRES      COMMERCIAL LAND 1995
BLMFG6    ....ACRES      MANUFACTURING LAND 1965
BLMFG9    ....ACRES      MANUFACTURING LAND 1995
BLRES6    ....ACRES      RESIDENTIAL LAND 1965
BLRES9    ....ACRES      RESIDENTIAL LAND 1995
BPOP6     ....PERSONS    POPULATION 1965
BPOP9     ....PERSONS    POPULATION 1995
BTA       ...._3"
                &
BTEMPA ....
BTEMPB....
BTEMPC ....
BTEMPD ....
BZZEL     ....TERS
BZZNAT    ....NATURAL FEATURES
BZZPOL    ....POLITICAL FEATURES
BZZRR     ....COMMUTER RAILROADS
BZZXWAY....EXPRESSWAY SKELETON
GC
B:_
```

Figure 5. Calculation on TEMPA-automobile ownership rate (core size for this version prevents storing more than two variables in core at once).

```
COMPUTE
B:TEMPA=POP6
***** THE CALCULATIONS HAVE BEN COMPLETED.  *****
** ANOTHER EQUATION ?    TYPE YES OR NO. **
B:YES
B:TEMPA=AUTO6/TEMPA
***** THE CALCULATIONS HAVE BEN COMPLETED.  *****
** ANOTHER EQUATION ?    TYPE YES OR NO. **
B:_
```

Figure 6. Displaying TEMPA as a function of its array location. Two extreme values indicate likely errors.

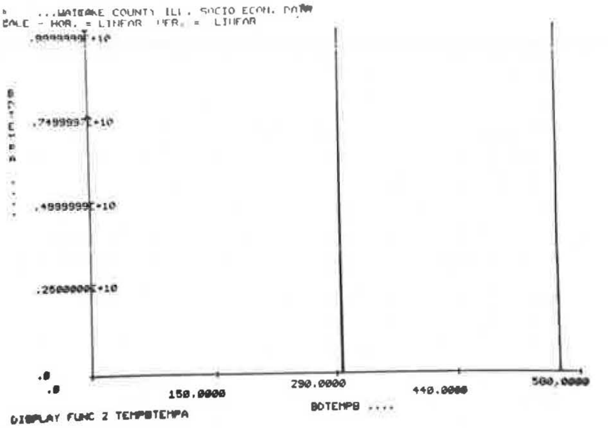


Figure 7. Editing the variables, to find and correct error in element 297 (making it 0).

```
EDIT 3 TEMPA POP6 AUTO6
EDIT
8:LIST 290 310
              TEMPA            POP6           AUTO6
     290      .0               .0             .0
     291      .30000           10.000         3.0000
     292      .30000           10.000         3.0000
     293      .43333           30.000         13.000
     294      .36667           30.000         11.000
     295      .45000           220.00         99.000
     296      .46667           60.000         28.000
     297      .10000E+11       .10000E-09     1.0000
     298      .40256           390.00         157.00
     299      .40536           1120.0         454.00
     300      .43333           60.000         26.000
     301      .32480           625.00         203.00
     302      .32500           40.000         13.000
     303      .30000           20.000         6.0000
     304      .32500           40.000         13.000
     305      .30000           20.000         6.0000
     306      .30000           20.000         6.0000
     307      .14000           100.00         14.000
     308      .12271           3390.0         416.00
     309      .14481           6160.0         892.00
     310      .0               .0             .0
EDIT
8:C 1 297 0
EDIT
8:L 560 570
              TEMPA            POP6           AUTO6
     560      .17500           720.00         126.00
     561      .45000           40.000         18.000
     562      .35000           20.000         7.0000
8:_
```

Figure 8. Map showing automobile ownership distribution in the region (overlayed with a grid in order to specify windowing).

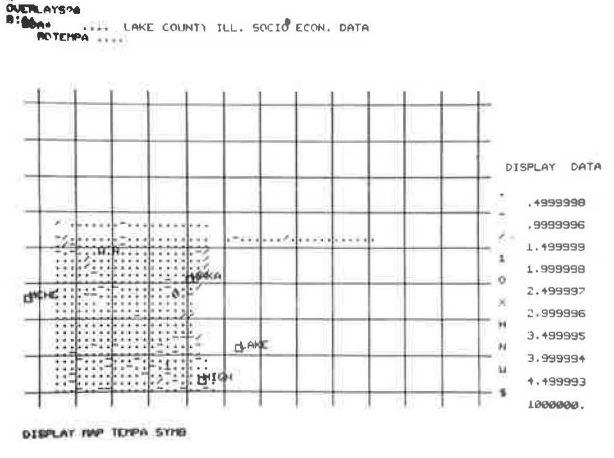


Figure 9. Directory of file BDSAMPLE (spatial data from a home interview, including a sample of 500 transit trips).

```
SHOW DIRECT
BDSAMPLE
BDSAMPLE
BDAIRLEN....MILE/100 AIRLINE TRIP LENGTH
BDARTIME....HOURS    ARRIVE TIME
BDDEST   ....RINGSEC DESTINATION OF TRIP
BDFPCOST....CENTS    FARE OR PARKING COST
BDINCOME....$ RANGE  TOTAL ANNUAL INCOME
BDLASTMD....         LAST MODE
BDNVEHIC....UNITS    NUMBER OF VEHICLES
BDORIG   ....RINGSEC ORIGIN OF TRIP
BDPRIOMD....PRIORITY MODE
BDPURFRO....         PURPOSE FROM
BDPURTO  ....        PURPOSE TO
BDSEQ    ....
BDSTTIME....HOURS    START TIME
BDTA     ....SEQUENTIAL NO.
BDTB     ....
BDTRPDUR....MINUTES  TRIP DURATION
BDTRVMOD....MODES    TRAVEL MODE
BDWKSTAT....STATUS   WORK STATUS
BDWTEND  ....MINUTES WALKING TIME--END OF TRIP
BDWTSTRT....MINUTES  WALKING TIME--START OF TRIP
GO
B:
```

Figure 10. Relative frequency distribution of trip starting time.

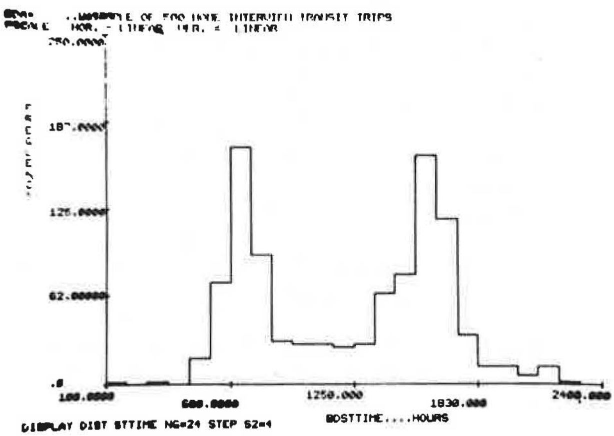
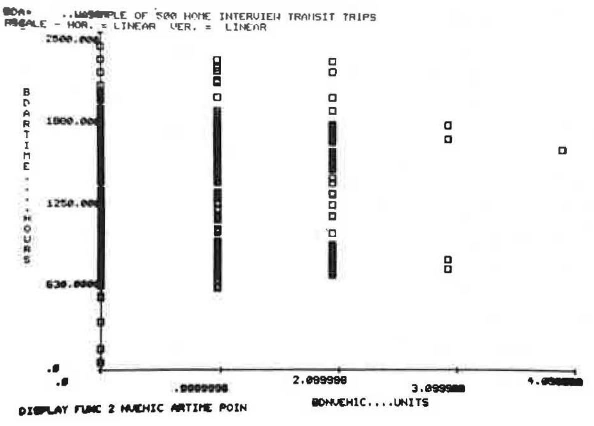


Figure 11. Arrival time distribution as a function of automobile ownership showing increased concentration of trips during rush hours with increased automobile ownership.



CRITERIA FOR BUS RAPID TRANSIT SYSTEMS IN URBAN CORRIDORS: SOME EXPERIMENTS WITH AN INTERACTIVE GRAPHIC DESIGN SYSTEM

Matthias H. Rapp*, Ecole Polytechnique Fédérale de Lausanne, Switzerland; and
Claus D. Gehner, Urban Systems Laboratory, University of Washington,

A man-computer interactive graphic system, developed at the University of Washington, is applied to a series of experiments designed to identify the characteristics of high-performance bus rapid transit systems for CBD-bound riders who reside in a generalized suburban corridor. The solution spaces of each of 12 problem cases, derived from the combination of several trip demand densities, highway network characteristics, and trip-making behaviors, are explored in an interactive graphic search process. The purpose of the search is to identify optimal solutions using two operator objectives, one maximizing profit and the other maximizing patronage within a given subsidy constraint. Furthermore, the search is aimed at identifying solutions that cannot be dominated in terms of any combination of patronage and profit, thus forming an envelope in the patronage versus profit-deficit space. From a comparison of these envelopes some general relationships between bus rapid transit operating characteristics and external conditions are developed. Similarly, relationships between system characteristics and external conditions are derived from a comparison of the penoptimal solutions. Finally it is shown that the penoptimal solutions may violate some of the environmental, social, or political constraints commonly imposed on public transit systems, and it is also demonstrated that, by complying with such constraints, the operator may incur significant losses.

•BUS RAPID TRANSIT (BRT) systems are a form of urban public transportation employing rubber-tired vehicles that operate for pickup and delivery on surface streets and for line-haul on limited-access highways or on exclusive rights-of-way. BRT systems may include parking facilities, allowing a user to drive to a park-and-ride lot as well as to walk to a bus stop on a pickup route.

A convincing case has been made stating that BRT systems have a high potential for successfully competing with the private automobile for trips that originate along an urban-suburban freeway corridor and that are destined to a major traffic attractor such as the CBD (1). It has also been shown that BRT systems have an economic advantage over rail rapid transit systems in corridors that generate less than 4,000 one-way peak-hour transit riders (2). BRT systems have been implemented and are being demonstrated in Washington, D. C., Los Angeles, and Seattle, and such systems are under consideration in many other U. S. cities.

It appears that existing or proposed BRT designs are based more on pragmatic considerations and qualitative judgment than on the results of analytical explorations of the great number of possible design alternatives. This may be due to the fact that the "urban transportation model system," the state-of-the-art tool for simulating urban

---

*Mr. Rapp was with the Urban Systems Laboratory, University of Washington, when this paper was written.

transportation system performance, is not sensitive to the entire range of design options of BRT systems (in particular the level-of-service options related to user convenience and comfort) and that the time and cost effort is too large to evaluate many different designs.

The lack of opportunity to broadly explore the space of BRT design possibilities may also be linked to the fact that there exists little theory about how BRT systems should be structured. Previous studies have been concerned with comparing BRT and rail rapid transit with respect to costs, speed, and capacity (2) and with operational problems of freeway metering and the use of exclusive bus lanes (3, 4). Many general questions concerning the network structure, level of service, and pricing of BRT systems are still unanswered. For example, there seems to be little knowledge concerning the general conditions under which BRT systems should provide area coverage, essentially whether every user should be offered the opportunity to walk to and from stations or whether service should be offered to only a few selected access points that are reached by the traveler by private automobile. We do not know how many such park-and-ride access points should generally be provided or whether they should be spatially distributed over the entire area to be served or only linearly along the corridor axis. We have little knowledge of how the relationships among level-of-service variables, transit patronage, and operator benefit behave under various external conditions.

This paper attempts to answer some of these questions about the design of BRT systems in urban freeway corridors by using a series of experiments of BRT design alternatives under various external problem conditions. The experiments employ a man-computer interactive graphic system for transit system design. Hence, as a second contribution, this paper demonstrates the application of the interactive graphic problem-solving methodology to answering a transportation research question.

OBJECTIVES

A CBD-oriented BRT system operating in an urban or suburban freeway corridor can be defined in terms of a set of design variables, i.e., the variables that are to be set in the design process. In a general case, the following variables may be included:

1. The number of park-and-ride lots;
2. For each park-and-ride lot, the location, size, and parking fee;
3. The number of transit routes;
4. For each transit route, the location of stops (i.e., the sequence of network links traversed), the number of transit vehicles, and the type of transit vehicles (i.e., passenger capacity, comfort level); and
5. The fare from each transit stop or park-and-ride lot to the destination.

A design process involves solving the complex problems of finding the optimal values for each design variable, where "optimal" usually refers to many, and sometimes conflicting, objectives. It is obvious that there is no general set of optimal design variables for BRT systems: A feature that is desirable in one case may not be efficient in another case. There are four principal external conditions that determine BRT system characteristics:

1. The amount and spatial distribution of travel demand;
2. The highway network characteristics, including the traffic conditions and CBD parking conditions;
3. The travel behavior of trip-makers (in particular their attitudes toward driving and riding transit); and
4. The objectives of and constraints on BRT operation (in particular, the extent to which economic profit is pursued versus the extent to which a high level of ridership is sought).

The objective of our study is to find some general relationships between external conditions on the one hand and BRT system features and their associated performance measures on the other.

Figure 1. Trip component impedances considered by modal split model.

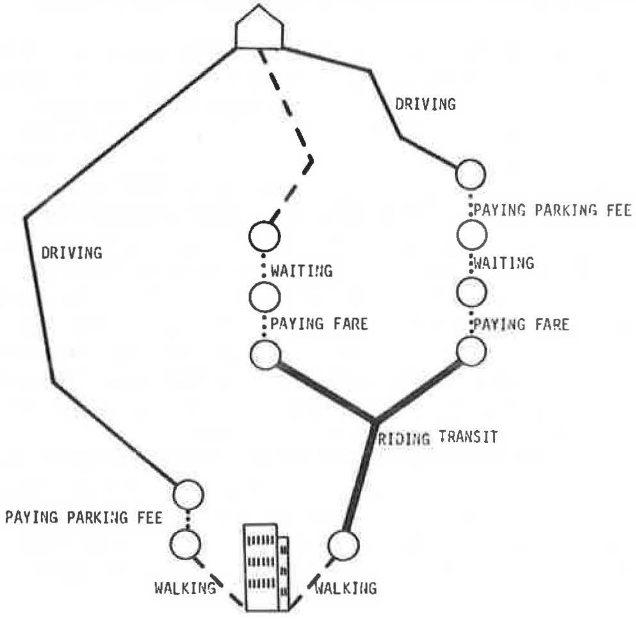
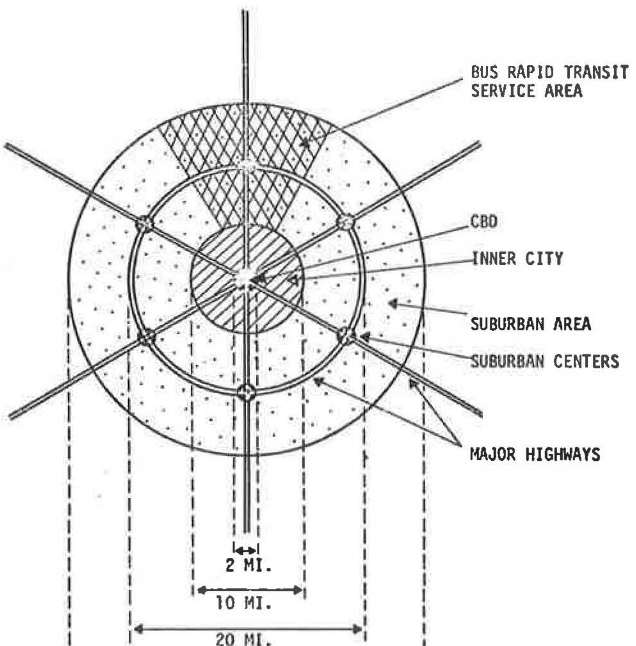


Figure 2. Hypothetical metropolitan structure.



EXPERIMENT DESIGN

To achieve the desired objective requires that a modal split-network equilibrium model be used to predict the performance of alternative BRT designs under varying external conditions. The modal split submodel partitions the total number of trips from each origin to their common destination into three sets: walk-and-ride trips, park-and-ride transit patrons, and drivers, i.e., nontransit users. The partitioning is performed on the basis of the sum of the perceived disutilities via each mode (Fig. 1) using a logit function. The transit trips generated by the modal split function are assigned on an all-or-nothing basis to the transit lines and parking lots on the paths with the minimum total impedance (i.e., perceived disutility) for each mode. A detailed description of the model is given elsewhere (5).

A hypothetical freeway corridor is used as the environment within which the BRT designs are generated. This corridor is assumed to be a sector of the metropolitan structure (Fig. 2). The street network within the corridor is shown in Figure 3. The population density is assumed to follow a distribution generated by the superposition of two negative exponential functions describing the influences of the CBD and a suburban center. The density surface for CBD-bound trips (Fig. 4) is generated by applying a gravity function to the population distribution.

To explore the relationships between alternative external conditions and the characteristics of BRT systems generated under these, we treat the former as variables. Specifically, we examine three trip density levels, which, when integrated over the entire BRT service area, correspond to total peak-hour demands, before modal split, of 5,000, 10,000, and 15,000 person-trips. Two cases of highway network speeds are considered. The first case represents the condition in which traffic flow on all links is fairly unrestricted and in which transit buses travel at the same speed as private cars. In the second case, the freeways and ramps are assumed to be congested; buses travel on exclusive bus lanes and are given preferential treatment at the freeway entrances. The travel speeds on all link types (Fig. 3) for the two cases are given in Table 1. Further, two alternative trip-making behaviors are assumed, each expressed by a set of average values for the disutility coefficients associated with the various trip components (Table 2). The first case is representative of a behavior that places a relatively high value on trip costs (including the hidden costs of operating an automobile) and a relatively low value on trip time and convenience. The second case is representative of the converse behavioral preferences. We shall refer to the former behavior as "money conscious" and the latter as "time conscious." The combination of these external conditions yields 12 solution spaces (three density levels x two network characteristics x two behavior patterns).

To be able to compare individual solutions within the solution spaces, we define two operator objective functions, toward which an optimization of the BRT designs are attempted. These objectives represent the private operator, who endeavors to maximize his profit, and the public agency, which seeks to provide maximum service subject to a subsidy constraint. In the case at hand, the maximum allowable subsidy is assumed to be $2,000 per day. This then defines the 24 problem cases under consideration here.

In addition to the variables mentioned, the model requires as input a number of other parameters, which are held constant throughout the experiments. Among these are land values, parameters of the park-and-ride lot cost model, and the average CBD parking fee for nontransit users, which is assumed to be $1.50 per day. Finally two alternative bus types with the characteristics listed below are assumed to be available; no constraint is placed on the number of such vehicles.

Type	Capacity: Seated	Capacity: Standing	Cost Per Hour (dollars)
1	20	5	12
2	40	10	16

Figure 3. Hypothetical street network within corridor.

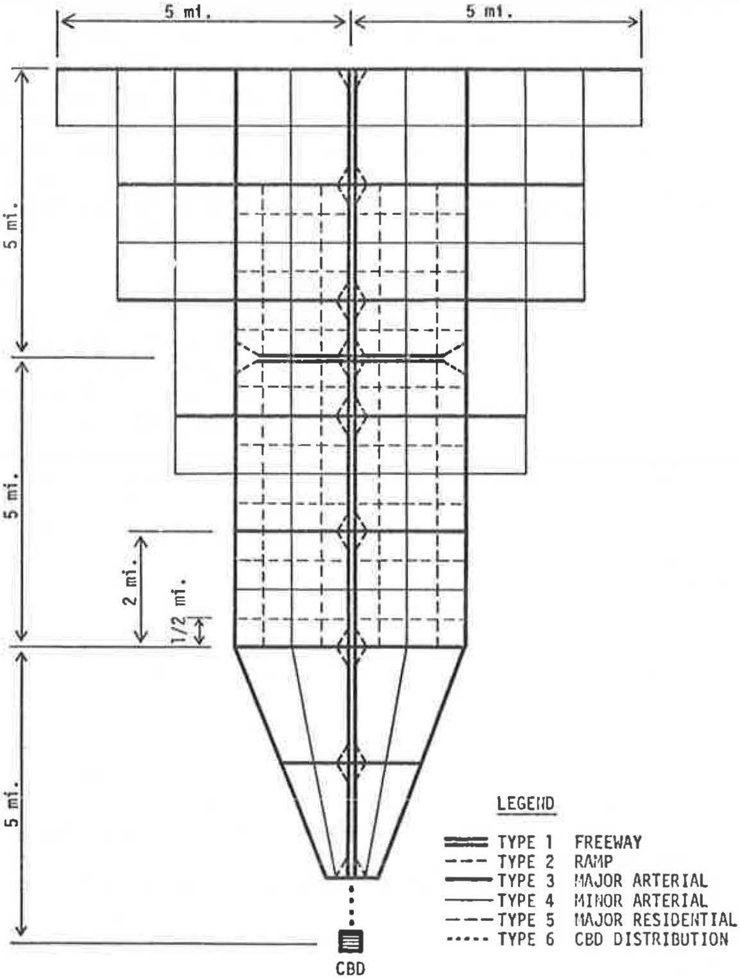
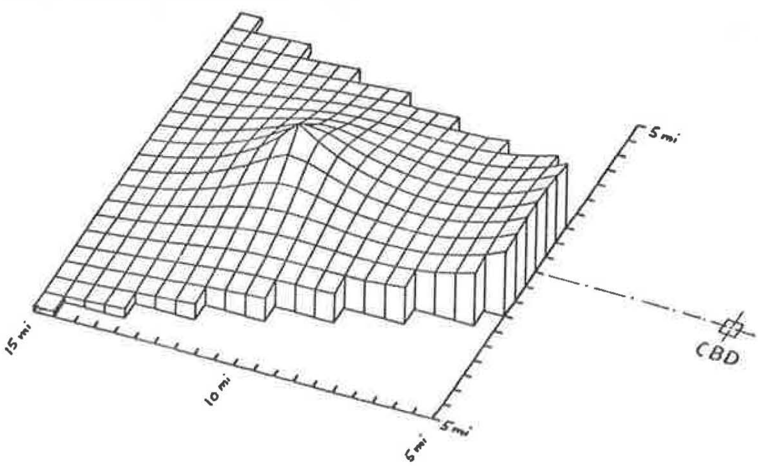


Figure 4. Hypothetical trip density surface.



In an attempt to explore the relationships between external conditions and BRT system features and their performance measures, two alternative approaches are possible. One approach would consist of analyzing the performance of a number of predetermined representative BRT designs under the 24 combinations of external conditions. The disadvantage of this approach is that it only allows a comparison of the relative merits of the predetermined solutions and those of the external conditions but does not allow general statements on the range of possible solutions, in particular the optimal among these. The other approach consists of a trial-and-error search that uses an implementation of the model in an interactive graphic environment. This approach is selected because it allows both the efficient exploration of the entire solution space for any given problem case and an approximation of the problem-specific optima.

RESULTS

An extensive man-computer interactive search was conducted by using the Interactive Graphic Transit Design System (IGTDS) (5, 6, 7), implemented at the Urban Systems Laboratory of the University of Washington. In IGTDS the user designs a BRT system by specifying routes, park-and-ride lots, vehicle characteristics, frequencies, and fares, and the computer immediately predicts and graphically displays on a CRT terminal patronage volumes, economic performance, and other impacts on his design. In this way 637 configurations, i.e., an average of 53 for each of the 12 solution spaces, are generated and evaluated.

The results of these experiments are presented in three parts. In the first part we present graphs of patronage versus profit-deficit for each of the 12 problem cases and analyze the effects of varying external conditions. We then show those configurations, among the ones tested, that best satisfy the two objectives for each of the 12 cases. These solutions are termed "penoptimal" in that an infinite number of designs to meet the experimental objectives are expected to add only imperceptibly to their optimality. From the correlation of the characteristics to the external conditions under which they have been generated, an attempt is made to suggest some general BRT design rules. Finally, some social and environmental constraints are imposed in addition to the subsidy limit on one of the designs to analyze and illustrate their effect on the penoptimal solution.

Before proceeding to the presentation of the results the reader should be aware of the following qualification: The results are a function of parameters entering the model as discussed in the previous section, which at this point represent only an estimate of the likely range of these parameters. Thus the analysis of results presented here emphasizes trends and relative comparisons rather than absolute values.

Trade-Off Between Modal Split and Profit-Deficit

For analysis of alternative BRT solutions, it is useful to view their performance in terms of the patronage attracted versus profit-deficit incurred, i.e., as one point in the modal split versus profit space. When all solutions are plotted in this way, it is found that there exists a set of solutions that cannot be dominated; i.e., no solution exists that has both a higher modal split and a greater profit (or lesser deficit). The locus of the points representing these solutions is called an envelope. This means that a set of external conditions imposes ultimate limits on the performance of a transit operator. It is further noted that the envelope is always convex with a negative slope that varies from $-\infty$ to 0. The envelope is very flat above a certain modal split, which implies that any further increase of service is highly uneconomical. A limiting modal split, less than 100 percent, is asymetrically approached by the envelope when it reaches a slope of 0. This is the ultimate modal split because the corresponding configuration is characterized by complete route coverage of the street network, zero fare, and zero headway, i.e., an infinite number of vehicles.

It is found that all solutions represented by points on the envelope have two common characteristics. First, the solutions represent demand-supply equilibria with all park-and-ride lot capacities being exactly equal to the respective park-and-ride demand and with the transit seating capacity exactly equaling the passenger flow demand in any

Table 1. Assumed network travel speeds.

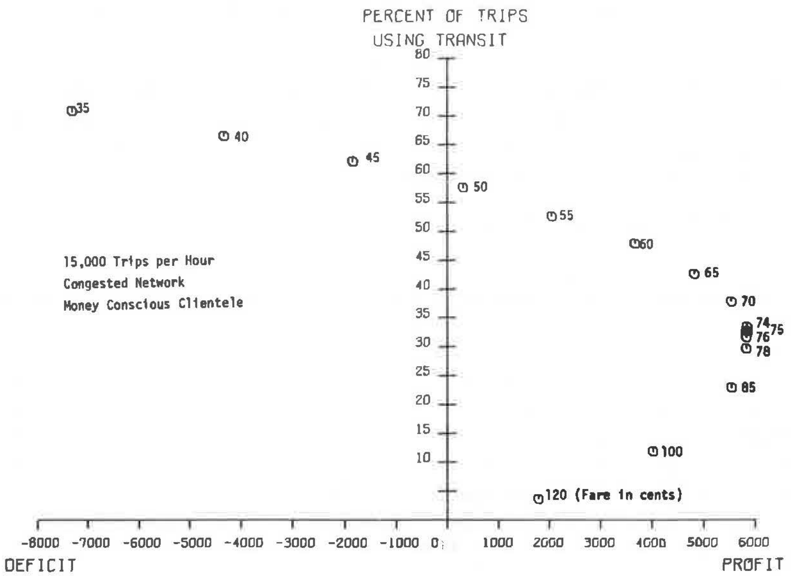
Case	Mode	Average Link Speeds (mph)					CBD Distribution Time[a] (min)
		Type 1	Type 2	Type 3	Type 4	Type 5	
Uncongested network	Bus	40	30	30	20	15	10
	Car	40	30	30	20	15	5
	Walk	—	—	3	3	3	—
Congested network	Bus	50	30	25	20	15	10
	Car	30	10[b]	25	20	15	10
	Walk	—	—	3	3	3	—

[a]Type 6.
[b]Represents an average delay of 2 min at freeway entrances and exits.

Table 2. Assumed impedance coefficients.

Activity	Case: Money Conscious	Case: Time Conscious
Walking	0.20/min	0.40/min
Driving	0.12/min	0.16/min
Riding		
Seated	0.08/min	0.16/min
Standing	0.24/min	0.48/min
Waiting	0.20/min	0.40/min
Paying		
Fare (one-way)	0.04/cent	0. 02/cent
Parking fee	0.02/cent	0.01/ cent

Figure 5. Average fares associated with solutions to problem 12.



transit line, i.e., neither standees nor empty seats. Second, solutions on the envelope are characterized by the absence of backtracking paths; i.e., there is no incentive for transit patrons to choose a longer path in order to save on fare or waiting time or both. It has been shown (5) that this condition is achieved if the combination of fares and headways at each stop produces a uniform impedance (e.g., lower headways compensated by higher fares and vice versa). Solutions with these two characteristics are called efficient solutions.

By investigating the solutions associated with points on the envelope it is also found that, the higher the modal split is, the lower the fare will be, as shown in Figure 5. It is further seen that, if the BRT system has been operating at a deficit, but efficiently, and a fare increase is instituted, then this fare increase, if accompanied by an appropriate reduction in service (i.e., again leading to an efficient service), will always result in a lower loss. (This reduction of deficit may not, in truth, be realizable because of labor practices prevalent in the transit industry.) If, on the other hand, an efficient BRT operation has been producing a profit, then an increase in fare (again accompanied by, for example, a decrease in the frequency of service) can lead to a higher or lower profit (Fig. 5). (Although possible in theory, this profit region of the envelope may never be reached in practice because of the political difficulty of raising fares and reducing service at the same time.)

The envelopes corresponding to the 12 problem cases may be compared by examining Figure 6, and the following trends can be noted.

1. As trip demand density increases, the envelopes shift to the right and upward. This effect can be attributed to economies of scale; for the same per-passenger cost, a higher density allows a lower headway or more extensive route coverage, both resulting in a higher level of service.

2. As congestion increases (assuming buses are given preferential treatment and exclusive rights-of-way), the envelopes shift to the right and upward, because transit becomes more competitive with the private automobile.

3. The envelopes corresponding to money-conscious behavior dominate those corresponding to time-conscious behavior throughout. This trend is expected in the situation of uncongested networks; however, one would expect time-conscious behavior to favor transit to a greater extent than money-conscious behavior in the situation of a congested network with busways because of the speed advantage of transit. This seeming paradox is explained by the fact that the saving resulting from not having to pay the CBD parking fee ($1.50) dominates and makes the money-conscious trip-makers more apt to use transit than their time-conscious counterparts. In this context, however, the attention of the reader is called to the percentage increase in profit at a given modal split level between the envelopes corresponding to time- and money-conscious behaviors. It is noted that this relative increase is greater in the case of congested networks than uncongested networks (e.g., the relative increase associated with problems 9 and 10 is greater than that associated with problems 11 and 12), which is a reflection of the expected preference for transit of the time-conscious behavior in the latter situation.

Characteristics of Penoptimal Solutions

The reader will recall that within each problem case two objectives were defined toward which optimization of the design was attempted, namely, maximizing profit on the one hand and maximizing modal split subject to a subsidy constraint of $2,000 per day on the other hand. The penoptimal solutions obtained in the interactive graphic search are shown for comparison in Figure 7. The following commonalities can be observed.

1. All park-and-ride lots are located at freeway entrances inasmuch as configurations with park-and-ride lots at other locations within the corridor were found to be inferior. This can be explained by noting that the freeway is the axis of symmetry for both the network and the demand distribution and thus represents the locus of lot locations that yield the highest accessibility.

Figure 6. Summary of modal split versus profit-deficit envelopes.

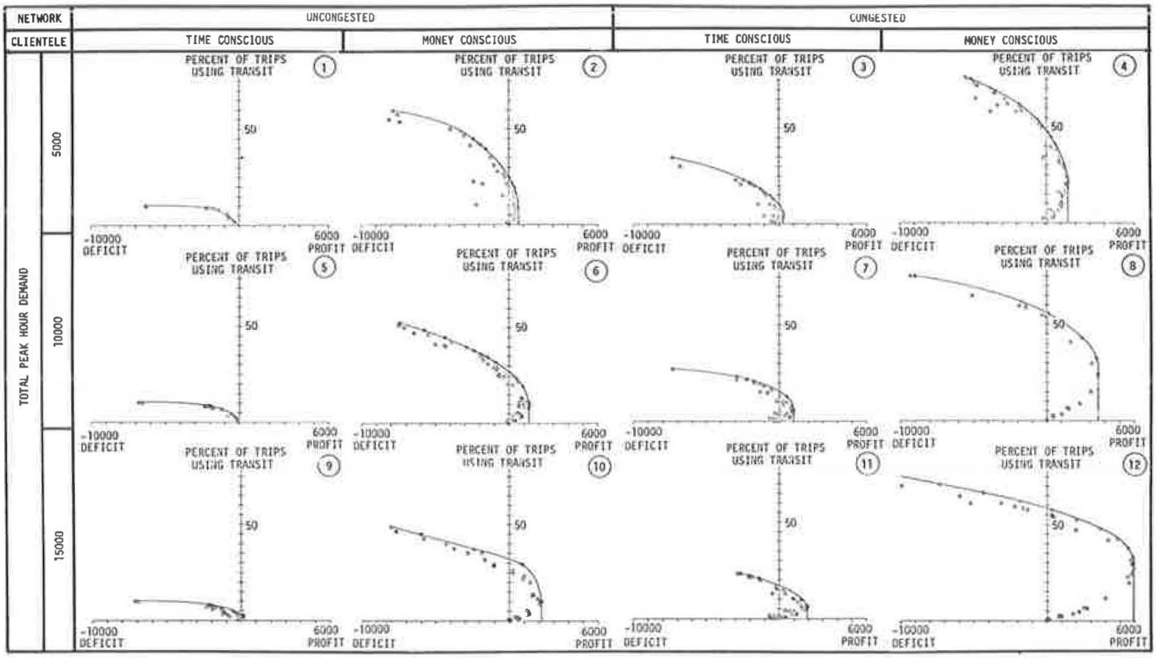
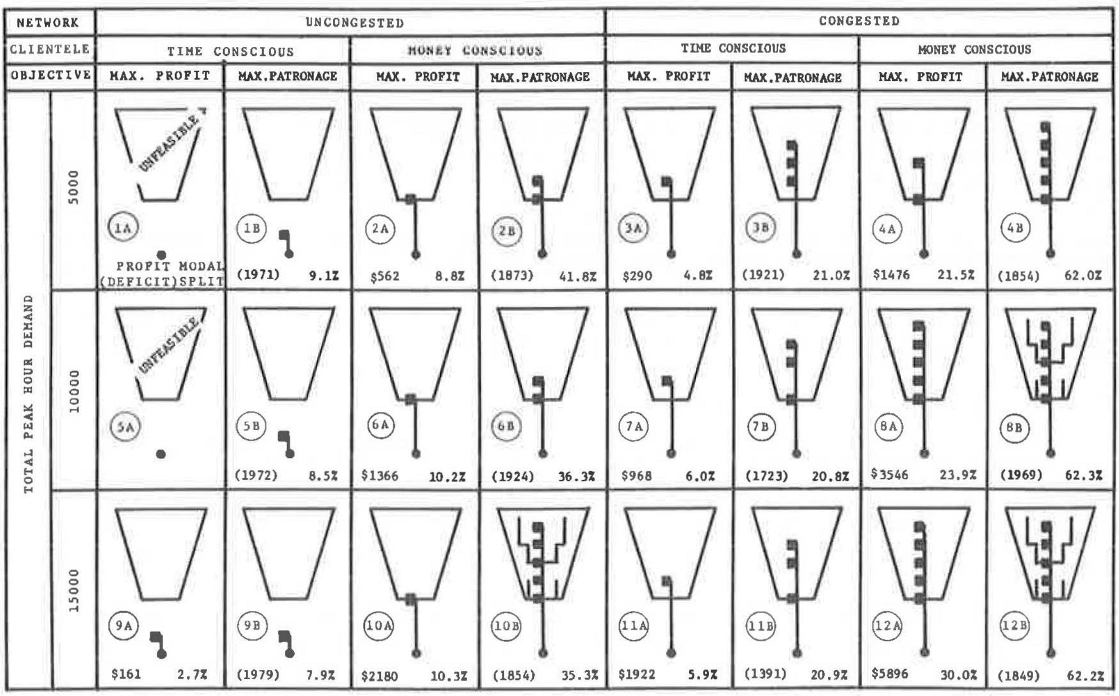


Figure 7. Summary of penoptimal solutions.



2. All park-and-ride lots are connected to the CBD by nonstop express routes. This can be rationalized by the fact that the time penalty associated with intermediate stops (dwell time and detour time) and the underused bus capacity on the outer part of the route outweighs the gain in frequency of service achieved by combining routes. In this context it appears that consolidating two or more park-and-ride lots that are infrequently served is more advantageous than merely combining the routes serving them.

3. As a result of the above and the apparent superiority of efficient solutions (capacity exactly equals the passenger flow demand) the frequency on lines serving park-and-ride lots is found to be proportional to the lot sizes. It thus follows that in order to provide decent headways park-and-ride lots must be quite large (e.g., to sustain a 5-min headway with 50-passenger buses, a lot must produce 600 trips per hour).

In addition to these commonalities, inspection of data shown in Figure 7 reveals certain trends from which some general BRT design rules can be derived.

It is found that most of the penoptimal solutions represent "pure" park-and-ride systems, i.e., configurations in which there exist only park-and-ride facilities and the routes connecting them with the CBD. Only in three cases were mixed park-and-ride/ walk-and-ride systems found to be superior. This situation occurs only when modal split maximization is the objective of the operator, and these cases are further characterized by a deficit, high trip densities, and a money-conscious clientele, the members of which are willing to walk to bus stops in order to save the costs of driving an automobile. However, even in these conditions favorable to transit, pure walk-and-ride solutions have been found to be inferior.

As density of trip demand increases, the extent of route coverage either is the same or increases. This is again a result of the economies of scale.

The modal split associated with maximum profit increases with increasing density, again because of economies of scale. However, it may be seen that the maximum modal split associated with the $2,000 per day subsidy does not increase with increasing density but remains constant or even decreases as in the sequence of penoptimal solutions 2B, 6B, and 10B. This seeming paradox can be explained by consideration of the fact that the per-person subsidy is higher in the low-density cases.

The solutions associated with the condition of the congested network show greater aerial coverage by their route networks than those associated with the uncongested network. Because of the speed advantage of transit in the former case, more patrons are attracted, who pay higher fares, which sustains a higher level of service and which in turn allows a more extensive route coverage. This speed advantage also provides an incentive to intercept park-and-ride transit users at a greater distance from the destination.

The solutions generated under the condition of money-conscious behavior show a greater route coverage than those associated with the time-conscious behavior. Again the high CBD parking fee is responsible for inducing more of the former to choose transit, which results in higher revenues and which sustains a higher level of service.

If the operator desires to maximize modal split, then any deficit constraint is always binding.

If the operator desires to maximize his profit, he should charge high fares and offer minimal service, except in those situations that highly favor transit usage (i.e., congested network, money-conscious behavior). It is understood that in the real world additional constraints that prevent the operator from attaining the maximum profit state are usually imposed.

In most cases it is found to be advantageous to use the larger of the two vehicle types. Only in the solutions associated with low modal splits and where the clientele is time conscious is use of the smaller vehicle more profitable, because of the critical influence of waiting time at transit stations.

Effect of Additional Design Constraints

The design of transit systems in our experiment was not subject to any environmental, social, or political constraints because we wanted to explore the unlimited solution spaces of the 12 problem cases. However, in any real-world design situations, such constraints are likely to be imposed. Therefore, in this section, we speculate about how some typical additional constraints would affect the optimal configurations we found and the conclusions we drew from these.

When the optimal transit designs generated for the 12 unconstrained problems are examined (a problem is "unconstrained" if it is subject to a maximum subsidy allowance but not to any additional constraints), the question arises whether any of the design variables or impact variables have values that would be clearly unacceptable to the transit operator or to the public at large. To answer this question, we selected a set of constraints that reflect the value ranges of transit system attributes that are sometimes mentioned in transit system planning. The following constraints are included:

1. No transit line should have a headway of more than 30 min;
2. No park-and-ride lot should be larger than 1,000 stalls;
3. The fare from any origin should not exceed 60 cents;
4. At least 20 percent of the clientele should live within a 5-min walk of a bus stop; and
5. At least 40 percent of the clientele should live within a 5-min drive of a park-and-ride lot.

The performance of the 24 optimal configurations with respect to the five constraints is shown in Figure 8. In each diagram the value of the performance measure has been normalized relative to the maximal allowed value. In Figure 8d and 8e, the constraints have been reversed in order to depict all constraints as "less-or-equal" inequalities. It can be seen that the constraints concerning park-and-ride lot size and percentage of trip-makers within a 5-min walk are in most cases violated. The largest violation occurs in the optimal solution to problem case 6, which has only two park-and-ride lots, one of them with 7,482 stalls. The fare constraint is violated by the solutions to most problem cases in which the operator attempts to maximize his profit. All solutions comply with the 30-min headway requirement.

To determine whether the optimal configurations can be modified to conform to the constraints, we undertook a new search on problem case 6, the solution that violated the constraints most. The search was again subject to maximizing total patronage subject to the maximum deficit allowance, in addition to the five new constraints. The best solution we found after nine trials is shown in Figure 9. The solution features five park-and-ride lots (as opposed to two lots) and four walk-and-ride transit routes. These routes are necessary to bring the required 20 percent of the riders within the 5-min walking perimeter, but they are all associated with deficits, i.e., the fare-box revenues on these lines do not cover the costs of the buses.

Of the six constraints, three are binding and three have a slack. The binding constraints are the maximum deficit allowance, the maximum parking lot size, and the minimum percentage of trip-makers within a 5-min walk to station. (The small amounts of slack associated with these variables could be reduced by conducting a more extensive search.) Not binding are the constraints concerning maximum fare (with a slack of 23 cents), maximum headway (with a slack of 14 min), and minimum percentage of trip-makers within 5-min drive time to lots (with a slack of 39 percent).

Most significant is the modal split associated with the solution. Whereas the unconstrained solution showed a modal split of 36.3 percent (Fig. 7, case 6B), the best constrained solution shows a modal split of 29.3 percent. In other words, the constraints "cost" the operator a loss of 7 percent patronage or, with 200 operating days, 140,000 trips per year. Similarly, if the patronage of the unconstrained problem were to be achieved, a significantly greater deficit would be incurred. This study is not the place for a philosophical discussion of whether the environmental, social, or political benefits associated with these or similar constraints justify their costs, but we have demonstrated that it is important to question the typical planning constraints by comparing the effectiveness of the optimal unconstrained solution with that of the optimal constrained solution. The fact that the man-computer interactive graphic methodology enables a user to do this is one of its most significant features.

Figure 8. Performance of optimal configurations given typical design constraints.

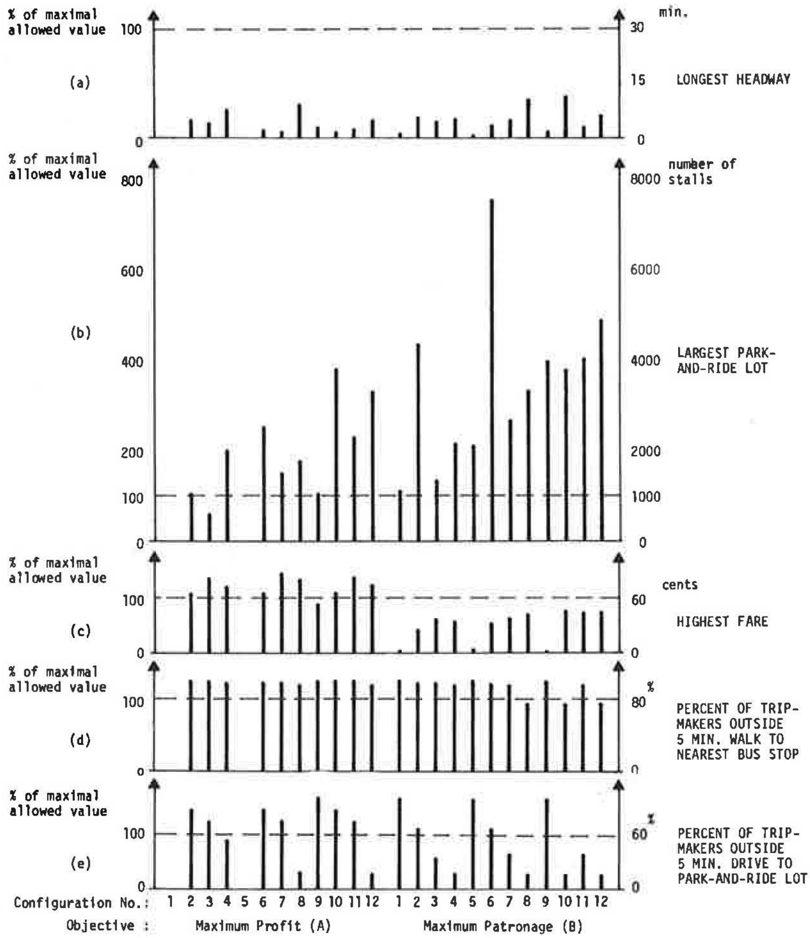
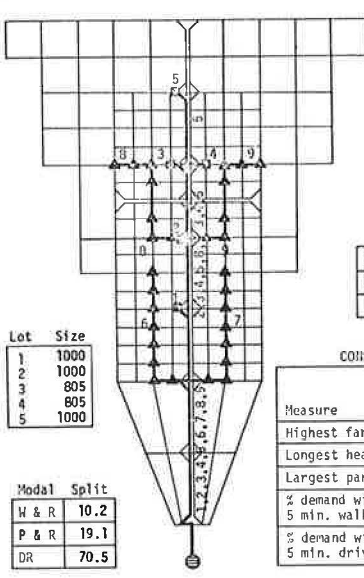


Figure 9. Best configuration for problem 6 subject to additional constraints.

DEMAND: 10,000 NETWORK: Uncongested BEHAVIOR: Money Conscious



Line	Veh. No.	Type	Headway	Fare
1	9	2	5.3	37
2	10	2	5.3	37
3	12	2	5.1	37
4	11	2	5.7	37
5	13	2	5.1	37
6	5	2	12.6	37
7	4	2	16.3	37
8	7	2	10.4	37
9	6	2	12.5	37

	Lot	Line	Total
Cost	1263	6160	
Rev.		5436	
Ben.			-1987

Lot	Size
1	1000
2	1000
3	805
4	805
5	1000

CONSTRAINTS

Measure	Max/min allowed value	Obtained value
Highest fare (cents)	60	37
Longest headway (min)	30	16.3
Largest parking lot	1000	1000
% demand within 5 min. walk time	20	23
% demand within 5 min. drive time	40	79

Modal	Split
W & R	10.2
P & R	19.1
DR	70.5

CONCLUSIONS

The results of our experiments suggest that, if the environment within which BRT systems operate changes from low to high trip demand densities, from uncongested to congested highways with a policy of preferential treatment of transit, and from an automobile-oriented society to one that is more transit oriented, the potential for a profitable BRT operation will increase substantially. However, under all but the most favorable conditions a profit can only be produced if high fares are levied and minimal service is offered, i.e., if transit caters only to a small share of its potential market; a low-priced, equitable, and extensive service is likely to incur a deficit. It has also been demonstrated that the per-passenger subsidy required to attract marginal patronage rises sharply above relatively low modal splits.

Moreover, it should be noted that the results reported here are optimistic vis-à-vis the situations experienced by actual transit operators in most urban areas for two reasons: First, our analysis focused exclusively on optimal configurations, whereas it is not likely that real-world transit designs, which have changed little in many decades, represent the optimal solutions under the prevailing conditions. Second, the BRT systems studied represent one of the most efficient types of bus transport service, namely, the many-to-one system. Consequently, the economics associated with many-to-many origin-destination service, which is the more common urban transport requirement, must be expected to be even less favorable to transit.

The interactive graphic problem-solving process used in these experiments has proved to be invaluable for efficient exploration of the respective solution spaces. Without it the results leading to the conclusions presented could not have been obtained. It is felt that the interactive graphic approach has been established as a useful tool in answering planning research questions.

REFERENCES

1. Wilbur Smith and Associates. The Potential for Bus Rapid Transit. Automobile Manufacturers Assn., Detroit, Mich., 1970.
2. Deen, T. B., and James, D. H. Relative Costs of Bus and Rail Transit Systems. Highway Research Record 293, 1969, pp. 33-53.
3. McDermott, J. M. Operational Effects of Automatic Ramp Control on Network Traffic. Highway Research Record 202, 1967.
4. Stover, V. G., and Glennon, J. C. Bus Rapid Transit on Urban Freeways Using Traffic Surveillance and Control. Traffic Quarterly, Vol. 23, No. 4, Oct. 1969.
5. Rapp, M. H. Planning Demand Adaptive Urban Public Transportation Systems: The Man-Computer Interactive Graphic Approach. Urban Transportation Program, Univ. of Washington, Seattle, Res. Rept. 71-4, 1972.
6. Rapp, M. H., Goldman, M., and Smallwood, R. The Interactive Graphic Transit Simulator, A Tool for Planning Node-Oriented Transit Systems. Urban Transportation Program, Univ. of Washington, Seattle, Res. Rept. 7, 1971.
7. Rapp, M. H. Transit System Planning: A Man-Computer Interactive Graphic Approach. Highway Research Record 415, 1972, pp. 49-61.

SPONSORSHIP OF THIS RECORD

GROUP 1—TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Charles V. Wootan, Texas A&M University, chairman

Task Force on Interactive Graphics
Yehuda Gur, University of Illinois at Chicago Circle, chairman
Edgar M. Horwood, Lloyd Pernela, Matthias H. Rapp, Earl R. Ruiter, Jerry B. Schneider, Martin Wachs

James A. Scott, Highway Research Board staff