

Physical planning for geographic areas such as cities, counties, or regions can be greatly simplified if the planner can display the entity under consideration together with land-use and socioeconomic data and can interact easily with that data to modify the presentation and redisplay it. Presented in this paper are a system and a language to aid such physical planning and user experience with the system and language. Further research on graphic presentation, the incorporation of models and statistical routines are also discussed.

The development of software systems to aid in physical planning

by B. S. Smedley

The terms *physical planning* or simply *planning* are used in this paper to refer to those activities concerned with planning the control and development of the social, economic, and physical systems of an area. Discussed here in particular are such activities as they are undertaken in the United Kingdom, where they are performed by the local government authorities.

Hitherto, the use made of computers in physical planning has been disappointing when compared with that made in other application areas. This paper considers reasons for this, by examining the nature of the physical planning process, and the implications and problems of designing a computer-based system to be used as a tool by those whose task it is to plan our physical environment.

The basis of this paper is a research project undertaken between 1972 and 1974 at the IBM United Kingdom Scientific Centre to investigate the role of information systems involving geographic (i.e., spatial) data bases as they relate to the work of local authority planning departments. This has resulted in the implementation of an Urban Management System.^{1,2} This paper describes the system design and the means used to evaluate the concepts incorporated therein—including their use by non-data-processing-oriented planning professionals. The penultimate section describes continuing research, and the last section draws some comparisons with related work elsewhere within IBM.

Copyright 1978 by International Business Machines Corporation. Copying is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract may be used without further permission in computer-based and other information-service systems. Permission to *republish* other excerpts should be obtained from the Editor.

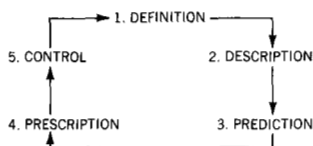
The emphasis throughout the work described here has been upon applications research rather than systems research. The overall objective has been to provide the planner with a set of tools with which he can tackle his work more effectively. Consequently, it is considered more important to evaluate its use in the application area, rather than the implementation itself. This topic is discussed further in the final section.

There has been a growing awareness of the relevance of the computer to applications in physical planning. Many local authorities have implemented information systems or are in the process of so doing, on the basis of property information relating to their area. One example is the LAMIS system of the Leeds Metropolitan District Council in the United Kingdom. The Department of the Environment has undertaken a number of studies and has issued several reports^{3,4} relevant to this area of study that act as the focal point for most discussions among practitioners from local authorities, research institutions, and other interested parties.

Whilst the organization of planning in the United Kingdom may differ from that in most other countries, the basic functions apply wherever and however planning is practiced. Thus, although our research has been specifically directed with the United Kingdom in mind, the techniques used have relevance worldwide. In general, the principles and concepts apply wherever professional but non-computer-oriented personnel need to have access to the power of the computer in order to analyze, manipulate, and process information that is relevant to their function.

Nature of physical planning

Figure 1 The physical planning cycle



The physical planning process is a five-stage cycle, as shown in Figure 1. The stages are the following:

- *Definition:* Agree upon the objectives for the plan.
- *Description:* Survey and analyze the existing situation, through data and information collection and processing.
- *Prediction:* Forecast the likely future situation that might result from specified actions or policies.
- *Prescription:* Evaluate the results of various policies against the agreed objectives, and select the most suitable strategy.
- *Control:* Implement the chosen strategy through control over development, and monitor actual progress against that planned in the strategy and the agreed objectives. This may lead to a change in the strategy implementation. A review of objectives and a revision of the plan may be called for if circumstances are not as originally anticipated, or if the basic assumptions prove to be false.

Table 1 Topics studied in physical planning

<i>Physical</i>	<i>Social</i>	<i>Economic</i>
Land use	Population	Employment
Housing	Education	Industry
Transport	Recreation	Shopping
Conservation	Community welfare	Income
Utilities (gas, water, etc.)		

The starting point for computer systems in planning has been the administration of development control. This is the processing of applications for development permission, from submission through approval or rejection to completion. Such systems are not part of the planning process, since the decision making is performed outside the system. A given decision, however, itself may be influenced by previous decisions recorded in the system.

The Urban Management System (UMS) is the outcome of research into the requirements of a computer system that is designed to assist in the second stage of the planning process, that is, to describe the existing situation. Table 1 shows some of the subject topics that need to be analyzed as part of this process. The diversity of the data base is readily apparent, as is the conceptual difficulty of classifying the data. This is a fundamental problem in planning, i.e., how to define the base. There are several other important factors to note.

For example, there are obvious interrelationships among such topics as population and employment, transportation and recreation. However, these relationships are very difficult to quantify. In many cases, the planning authority does not maintain the operational data with which to construct the subject data base, but either has to rely upon another agency to provide the data (e.g., central government) or else must undertake a survey to obtain it. In the United Kingdom, income data are almost impossible to obtain in detailed form, yet must be taken into account by the majority, if not all, of the plans. Much of these data need to be analyzed in their distribution over both time and physical space. Time series data are not new to computer systems, but spatial data are, insofar as information systems are concerned. Spatial data present problems in collection, representation, and storage. These factors are considered in subsequent sections of this paper, which discuss system concepts, the methodology used to implement these concepts in the UMS prototype, and their evaluation in the data analysis phase of the planning cycle.

Consideration of the relevance of computer systems to the other stages of the planning cycle, which is the subject of further ongoing research, is given at the end of this paper.

Implications for system design

Incorporation of the requirements for a data analysis system in physical planning, as discussed above, gives rise to the following implications for system design:

- Data handling
- Data operations
- Treatment of spatial information
- Direct use by professional planners
- On-line unstructured operation

data handling

Consider first the requirement for handling large amounts of data in differing formats. This is not a requirement for a data base system per se, since the transaction processing applications are not part of the main planning function. Rather, a data analysis system is required, but one in which it is possible to access data of widely differing formats and at different levels of aggregation. At the same time, the data analysis system should be of such a nature that the structure of the data contents in no way limits access to any part of the data base.

operations on data

Second, the operations to be performed on the data include both access to single items or records, and the manipulation of sets of data by single operations. Furthermore, it must be possible for the user to create new sets of data dynamically in any required shape or form, and then it must be possible to treat these data sets in identical fashion to those already existing in the data base. Thus, for example, it should be possible to determine the basic population statistics (i.e., numbers in age groups, males and females, single and married, etc.) for any specified area for which such statistics are collected. Alternatively, the planner may wish to identify those areas in which certain conditions exist (e.g., more single females than males) and treat this as a new set of information. The planner may then wish to determine the basic statistics for these areas alone. He may also wish to create new sets of data based upon two or more of the original topics. For example, he may combine population and housing data by area, and calculate a frequency distribution of household size or the proportion living in substandard accommodation.

spatial information

The spatial nature of the information to be processed in a physical planning project could be handled by considering it to be a collection of character or numeric data that define location in certain

terms, e.g., postal address, postcode, physical area identifier (e.g., zone 26), or grid square reference, and the processing of such ordinary character or numeric information. Alternatively, a segment referencing system, such as the Dual Independent Map Encoding (DIME) method of the U.S. Bureau of the Census, could be adopted, in which the spatial information is used as an index to the basic information.

Another way, and the one adopted in the Urban Management System, is to consider spatial information as simply another data type. Thus the system, as well as handling character and numeric data, must handle locational data, and as far as logically possible perform operations equally on all three types. The requirements for physical planning are to manipulate spatial values that specify points, lines, and areas (e.g., house location, route of road, boundary of a district). The General Information System for Planning (GISP) report, which acts as a focus for United Kingdom planning implementation, proposes a locational-referencing scheme based on two-dimensional national grid coordinates. This scheme is followed in the Urban Management System. Locational referencing offers several advantages over the alternative methods described above. It is more flexible, caters to the irregular shapes that abound in the real world, and enables the aggregation of data spatially and the calculation of distances, surface areas, etc.

By treating spatial information as another data value, such information can be regarded by the user as just as basic a part of the data base as any other information. This increases the system's capabilities by extending the scope of data analysis. Thus not only can all houses with more than three bedrooms be identified, but also so can all houses in any specified area—as long as the location of every house and the boundary of that area are recorded.

The provision of a further data type necessitates additional system functions, in order to exploit the extra capacity of the system. Such functions should be considered as natural extensions to those functions already provided for the manipulation of character and numeric data. Table 2 illustrates this point with some examples.

The remaining two implications for system design revolve around the direct use of a system by a professional planner. The underlying principle throughout our work has been the necessity for the planner to interact directly with the system himself. Interactivity is necessary because of the unstructured nature of planning work and the near-certainty that the data available to the planner are incomplete, that they come from a variety of sources, and that

**use by
planners**

Table 2 Examples of functions

<i>Type</i>	<i>Numeric, character</i>	<i>Spatial</i>
Logical	$A > B$	Point in polygon, line crossing, area overlap
Conversion	Character $\leftrightarrow$ numeric	Spatial $\rightarrow$ numeric (e.g., length of line; area of surface)
Arithmetic	$A + B$	Union of areas; difference between lines

they exist at different and incompatible levels of detail. The need is therefore to cater to problem analysis as well as to problem solving.

Minimal restrictions should be placed upon a user with regard to any required data-processing knowledge, and he should not need to rely on the communication of his needs via a systems analyst to a programmer in order to solve his planning problems. Instead, the planner needs to explore the data that are available in an unstructured way. That is, the planner may browse and not be limited to any predefined operations, as would be the case if batch application programs were provided.

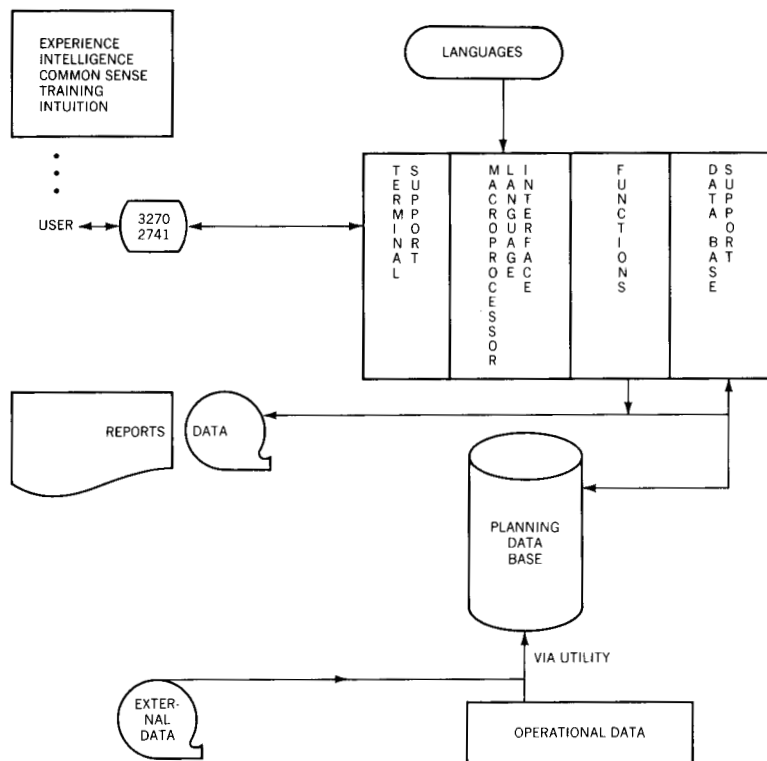
On-line usage is essential because the planner can formulate his operations in the light of the data he sees and as he interprets the implications of the results of the previous operations. This further implies a language facility that can grow with the user or change with different users, but yet be as transparent as possible to each user. Thus neither the language nor the system distracts the planner as he seeks to understand the relationships that reflect the environment for which he is planning.

System methodology

The background to the implementation of the Urban Management System has been described. This section considers factors that affect system design, as they were discovered during our research.

An overview of the system and its mode of use is shown in Figure 2. The terminal support has been made as simple as possible. A simple keyboard entry approach uses either an IBM 3277 display unit or an IBM 2741 typewriter terminal. Only a single user terminal is supported by the prototype because it is not the purpose of the project to investigate the problems of multiuser access.

Figure 2 Overview of the Urban Management System



The approach taken in the provision of the user interface is to opt for a macroprocessor (MP/3).⁵ By using a macroprocessor it is possible to define a number of languages, all of which can be translated into the common internal system language. MP/3 gives the additional facility of being able to define new macros dynamically. Thus, not only can the system be altered by linking to a different language definition, but also a user can modify his view of the system during a session. Any such modifications would obviously require the user to have an understanding of the macroprocessor itself and, depending upon the modification, to have an understanding of the Urban Management System. This ability is an important factor in giving the prototype the open-endedness and flexibility necessary to evaluate the needs of potential users.

Initially, two user languages were created. The first was a verbose and cumbersome English-like language that was easy to understand, once the keywords had been grasped. A number of synonyms were provided so that the user did not have to remember particular keywords, and the system would prompt the user for additional information in certain situations. This approach was the first to be chosen, and it proved to be suitable for those seek-

languages

ing an introduction to the system, or who expect to use the system only occasionally. Such users wanted as little training as possible. For continued use, however, the first language was not suitable. Consequently, the second language was designed to be more functional and concise. A more mathematical or APL-like approach that relied heavily upon a symbolic notation was adopted. The limitation of the keyboard characters available restricted the symbolic set so that in certain cases composite symbols were used.

The following are statements in both languages to perform the same task.

Language 1

```
CHOOSE HOUSES  
  
SELECT ENTRIES WHERE BEDROOMS>3  
  
DISPLAY OWNER, STREET, BEDROOMS
```

Language 2

```
*T(OWNER, STREET, BEDROOMS) ←HOUSES [S:BEDROOMS>3]
```

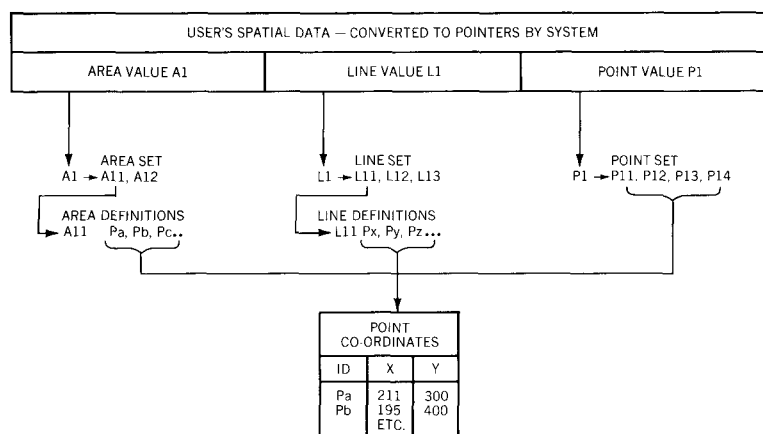
The terminal support plus the macroprocessor-based language interface enables the professional planner to use the system and have direct interactive access to his data with minimal data processing knowledge. In this way, the Urban Management System enables the interplay between the quantitative data held in the system data base and the distributed (qualitative, subjective) data base represented by the experience, intelligence, professional training, and intuition of the user. Only by enabling the two data bases to interact can planning problems be resolved effectively.

functions

The majority of the programming effort has been concentrated in the functional part of the system. Once an initial functional capability was agreed upon, it remained to decide those functions that could be handled by the data base software and those to be implemented directly. All programming used the PL/I language, each function being implemented as a module with calls to subroutines as required. Thus a selection function could be implemented by a single call to a data base routine. An aggregation function would be written as a series of calls to the data base read routine, with returns each time to the higher level to compare data and aggregate as necessary.

The implementation of the geographic data type occurs in the functional section. Point, line, and area values are supported in such a manner that complex structures of each sort can be treated as single values. For example, a river and all its tributaries can be a single line value. Similarly, a single area value can comprise a set of unique areas (not necessarily connected as in a network).

Figure 3 Representation of spatial data in the Urban Management System



For example, some of the areas may be regarded as "holes," and the total may represent all the parkland in a county.

Implementation at the point-line-area level is necessary because the data base support is only for numeric and character data. Consequently, geographic values have to be mapped into strings of coordinates that can then be processed as numbers or character strings by the data base software. This has been achieved using a system of pointers as shown in Figure 3.

In order to present geographic data to the user as a single value, functions that operate on geographic values also have to be implemented in the function section. This is accomplished by receiving strings of numbers from the data base and interpreting them as geographic values. Alternatively, the system can receive geographic values from the user and reverse the above process. In addition, processing functions can be invoked to perform such operations as point-in-polygon searches, line-crossing tests, calculation of physical values (e.g., lengths or areas), and other geographic functions that might be identified.

The data base software chosen to fulfill the data base support requirement of the system is the relational data base software implemented at the United Kingdom Scientific Centre and known as the Peterlee Relational Test Vehicle (PRTV).⁶

**relational
data base
support**

The design of the Urban Management System (UMS) is such that all access to the data base software is through a single internal interface. Thus, if desired, an alternative data base system could be utilized. Considering the requirements of the system, however, the relational data base is believed to best offer the required flexibility and functional capability. Basically, UMS relies heavily

on the ability of PRTV to manipulate blocks (or files) of data; to perform set operations (union, intersection, difference), selection, and read and write operations on these blocks; and to join them together according to data values in each block. These functions could be provided at a higher level, i.e., in UMS itself, but at the considerable expense of loss of efficiency and performance. In addition, PRTV has no concept of key fields. Consequently, data can be retrieved by reference to any field. The storage of data is left to PRTV, and is of concern neither to the user nor to the rest of UMS, apart from some conventions that have to be observed.

As with MP/3, there is considerable advantage in implementing the support required by UMS in existing, proven software. Furthermore, because PRTV had been developed and is maintained at the United Kingdom Scientific Centre, support and advice are close at hand.

Evaluation projects

Evaluation of the concepts embodied in the UMS prototype has been achieved in two diverse ways. First, a joint project was established with the Greater London Council in which a practical data base was created and interrogated, using UMS over a period of ten months. The second evaluation involved a senior planner from a local authority, who joined the Scientific Centre on a fellowship, and who worked full-time for seven months, using UMS to study a typical complex planning problem.

The main purpose of both evaluations was to examine the concepts proposed by our research in the light of their exposure to practical planning problems. In this way, concept feasibility could be established and also any deficiencies made known. The experimental nature of the system was stressed throughout, although its performance became of increasing significance, especially to the users.

data base

The data used in the Greater London Council evaluation related to land use and planning decisions. The land use data had been created as a result of the 1971 Greater London Land Use Survey, which recorded information for every occurrence of the use to which land, including property, had been put. The nature of the information recorded included street and property references, location (in national grid coordinates), floor space, type of use, employment data, car parking space, and a series of numeric codes that indicate spatial information, e.g., traffic zone and enumeration district numbers. In the evaluation, data for two of the thirty-three London boroughs were loaded into UMS.

Table 3 Summary of data used in the practical data base evaluation

<i>Land use data</i>	<i>Number of records</i>	<i>Fields per record</i>
Employment	46,560	23
Housing	30,199	17
Planning areas	35,488	15
Public buildings	1,490	14
Properties	35,541	39
Streets	2,199	7
Use	76,547	13
Total	228,034	(15,715,820 bytes)
<i>Planning decisions</i>		
Outstanding permissions	23,931	150
Current completions	30,485	150
Reserve completions	17,723	150
Expired permissions	1,478	150
Total	73,617	(31,287,225 bytes)

The second data set loaded was of planning decisions relating to individual developments within the Greater London area. These were of a strategic nature and of concern to the Greater London Council. The information recorded included total and net increase in floor space and parking space by use, date of permission and completion, and location. Table 3 summarizes the data used in the practical data base evaluation.

The procedure in the evaluation was for Greater London Council planners to use UMS and to access the data in order to fulfill their daily role as providers of strategic information to the planning and other departments of the Greater London Council, to central government, the London Boroughs, and other interested parties. The evaluation was running in parallel with the Greater London Council's own batch application system, which provided the same service. The evaluation lasted for ten months, during which time 54 terminal sessions were held, and occupied an elapsed time of 80 hours in total. A typical terminal session would last for two hours. This was found to be the maximum length of time it was reasonable to expect a user to interact with the system, and also enable useful work to be achieved. (Session time is a function of the overall loading of the host machine as well.)

In the second evaluation, a planning data base had to be created without resort to a local authority. This was achieved by purchasing the 1971 Census of Population data for a certain area and digitizing some large-scale maps that relate to the same area (Teeside). Thus land-use information was acquired. The objective was

**planning
problems**

Table 4 Summary of data used in the typical planning problem evaluation of the system

Population census—small area statistics for central part of Teesside (125 enumeration districts)		
Population:	125 records each with 471 fields	
Housing:	125 records each with 449 fields	
10% sample:	125 records each with 368 fields	
Total:	483,000 bytes	
Digitized map data covering same area (4 km square)		
Areas:	enumeration district boundaries	125
	physical areas	109
Lines:	Roads	622
	Railways	21
	Rivers	5
Points:	Houses	4,868
	Other properties	457
(Plus associated numeric and character information, e.g., property numbers, street names, use codes)		
Total:	approximately 1 million bytes in all.	

to investigate a typical planning problem—the provision of housing—using UMS both to ascertain the existing population and housing situation, and, by projection into the future, to consider the implications of housing policies. Table 4 summarizes the data involved in studying the given planning problem.

The two evaluations—data base and planning problems—covered a wide spectrum of planning activities, from the detailed data analysis in response to specific queries as in the case of the Greater London Council, to the strategic consideration of wider issues. Whereas with the Greater London Council the data existed to satisfy completely the problem at hand, the internal evaluation suffered slightly because all implications could not be considered. This was because it was impossible in the context of the work at the Scientific Centre to obtain all the relevant data.

In both evaluations, after a short and informal training period, the user was left to explore the system and data in order to gain more experience. He was then encouraged to use the system to help solve particular problems as they arose, or—in the second evaluation—as they were foreseen.

In neither evaluation did the design team attempt to impose any patterns upon the users, apart from those necessitated by the user-interface language provided in UMS. Since the emphasis in such information-oriented applications is very different from that in more scientific areas, it was believed to be important not to try to define very limited but controlled experiments. Rather, the objective was to observe what happened while such a system was being used by typical professionals. Conversely, no detailed anal-

ysis was made of the actual sequence of user-language statements entered because these would obviously depend upon the particular query being processed and the nature of the data involved.

Observations could be and were made, however, about frequency of use, circumstance of the situation, and other general matters, including conceptual ease or difficulty in operation. Much store was set by users' opinions and comments regarding particular features of the language and the system, although these had to be modified in the light of the designers' knowledge of the experimental nature of the implementation and its known deficiencies and inefficiencies.

Results of the evaluations

The joint project with the Greater London Council to evaluate the use of UMS in a practical situation has been made the subject of a separate report.⁷ The conclusions from that report and from the planning evaluation are drawn together in this section.

The evaluation projects confirmed both the feasibility of the UMS concepts and the capability of the implementation, by undertaking practical queries upon actual data. The amount of data held in the data base was sufficient to illustrate the capacity of UMS to handle and process queries against a large-scale data base in an on-line mode (40 million bytes of basic data in the Greater London Council study).

Of equal significance were the findings with respect to the direct use of a complex system by personnel with little or no training in data processing, and certainly no data base experience. The findings can be classified as follows:

- Understanding of functional capability
- Optimal arrangement of data
- Language basis of the interface
- On-line usage implications
- Training

For one to make the best use of a system, it is necessary to understand fully its functional capabilities. Only then is it possible to process queries against a data base in the knowledge that the results will be as logically correct as the contents of the data base allow. The user is then not misled by misuse of the system arising from insufficient or erroneous understanding. Besides understanding the functions available in the system, he must also appreciate the underlying concepts. UMS removes from the user the need to understand how the data are stored. However, in doing this, certain conventions are adopted, and the user must be made

**functional
capability**

aware of them. Otherwise, unexpected results can occur, affecting the credibility of the system.

**data
structure**

The ability to manipulate or perform sophisticated operations on large amounts of data can present performance problems. These can be minimized if the data are held in the manner most suited to those particular operations that are being performed. Obviously, this data arrangement may well vary according to the particular operation, and few users wish to be concerned about this aspect of their system. This matter can be eased by showing the user that performance improves if all redundant data are removed prior to each operation or series of operations. UMS requires data to be loaded into blocks, since the block is the basic collection of data that can be manipulated. Without the privilege of experience, it is natural to create these blocks to conform as nearly as possible to the way the original data are organized. Consequently, for any particular operation, much redundant data can be involved. Within a UMS environment it is better to use the UMS data structuring facilities to create new blocks and retain only those entries and attributes required for a series of operations, then revert to the original block for the next series. Again, some insight and experience are required to exploit this approach fully.

In practical terms this rearrangement of blocks can be achieved by formulating queries prior to an on-line session. Then, if necessary, a batch session may be run first to achieve the desired restructuring of the data, without having to wait at a terminal for this to be done. As an alternative and more acceptable approach, further research at the Scientific Centre has been investigating optimization within the relational data base context to discover how a system can better perform operations requested by a user by deferring and modifying them in the light of subsequent requests based upon the system's own knowledge of the data and its most efficient operations.⁸

**language and
interface**

Originally, one objective of the UMS project was to develop the ideal planning language that could be all things to all men. Because of the diverse natures of users, however, and the equally diverse applications that can be developed using UMS, it soon became apparent that this approach was impractical if not impossible. UMS has the ability to create a new language or macros based upon existing ones. Indeed, one person from the Greater London Council proved to be adept at enhancing his version of the language interface, once he had grasped a few fundamentals concerning the language processor. These enhancements were made to avoid much repetition of typing, where a series of operations involving different block names and/or attribute names had to be performed. Thus a higher level language was constructed with savings of the order of 90:1 on input in certain circumstances.

This ability to invoke a series of operations from one statement coupled with a parametric capability to vary the exact operation each time gives the user a simple programming capability. It was instructive to note how easily such a new concept could be grasped by a user, although admittedly he had some programming background. However, the results were encouraging enough to suggest that other users can benefit in this way. Hence the conclusion that, rather than trying to seek the perfect language, it is better to have a flexible interface and so let a series of languages be developed by users, perhaps with some guidance from a system expert.

No difficulty was encountered with the symbolic APL-like language. One user stated that after one hour's reading of the manual and one hour's practice, he was only limited in his use of the system by his own lack of knowledge about the data.

Experience in the evaluation projects has shown that true on-line interaction with complex systems is not yet a straightforward matter. The reason for this is the necessity to fulfill the following requirements:

on-line
usage

- Remember names and the coding conventions of the data contents.
- Plan one's operations in advance, and consider the best method if alternative approaches are possible.
- Consider what can be achieved in a terminal session and what can be performed better in a prior batch session.
- Consider the terminal device available in relation to prompting, display, and reference to previous operations.

The advantages of interaction can be nullified if the user does not appreciate performance criteria and expects instant responses to all his operations. (This issue becomes even more clouded if the system operates in a time-sharing environment.) There is a conflict between simply-entered commands that perform complicated operations and outwardly similar commands that are simple in operation. Only by experience—either direct or learned—can the user grow to understand this system performance conflict.

The system evaluation projects show that the best ways for users to become educated are the following:

training

1. A short formal period of study of theory and concepts.
2. Education in the use of the terminal.
3. Practical terminal sessions.

The first two can be completed in a few days, and then the prospective user is free to explore the system for himself. It helps greatly if the user is familiar with the data held in the data base,

and if a small subset of the data is made available for training so that response times can be shortened. When the newly trained user has access to the large data base he experiences greater difficulty. Considerations have to be made regarding performance, expected response, and the best methods for solving a given problem. The implications for a practical installation of any system like UMS are that while the system removes from the average user the need to understand any more than his data and how to solve his problem, each installation should have a central pool of expertise to load the data initially, consider the best format for it in UMS blocks, and to advise users regarding performance and method.

As to the type of person who can use UMS, the evaluation projects have demonstrated that those without any data processing knowledge can be as proficient as any others. This is held to be one of the key results of the evaluations. However, from a human factors viewpoint, it was observed that of two professional users, both very similar in age (young), qualification (degree level), and experience (several years in local authority planning departments), one was much more at home with the system than the other. This did not seem to be a deficiency of the system, its language, or the terminal, but rather a difference in approach to the same problem by two otherwise similar people.

**geographical
data**

In both the evaluations, spatial data were held in the data bases. However, such data were not used in any significant extent in the processes performed. During discussions with senior planners at the Greater London Council, the need for a geographic capability became evident. The conclusion to be drawn is that since existing information systems cannot handle geographic data, except by splitting grid coordinates into individual *X*- and *Y*-values, alternative means have been adopted to gain the required capability. This is shown in the data used in both evaluations by the existence of numerically coded fields that indicate geographic areas (e.g., 26 denotes Traffic Zone 26, with the spatial value of the zone not being recorded). Geographic selections can then be performed numerically, so long as the boundaries of such areas remain constant. When they change, much manual effort is required to update the records. If, however, the spatial value can be held, as is possible with UMS, this problem is removed.

Another reason put forward for the lack of geographic queries encountered in the evaluation projects was that the queries asked were at the strategic level and were thus concerned with the whole data, and not at any intermediate level that would have required at least selection of the data on a geographic basis.

Many local authorities in the United Kingdom are in the process of collecting spatial information relating to land use, property,

and population. They are also collecting planning control data. The intention is that such information can be used in several ways, such as the following:

- To enable graphic output of planning information, by means of plotters or line printer maps.
- To enable cross-referencing between otherwise incompatible data (e.g., between property and area information) on a flexible basis, and not to have to rely on manually entered references.
- As information in its own right for calculations, or to enable basic information to be aggregated according to a spatial criterion.

UMS provides the facilities for the second and third of these directly, and can act as the basis for the first by enabling the data to be stored. The system could then be the interface to a plotting or mapping routine.

The conclusions drawn regarding the use of geographic data are that spatial processing in physical planning, at least in the United Kingdom, is still in its infancy, but that the benefits to be gained are such as to encourage planning authorities to tackle the initial task of collecting the information necessary to provide the basic spatial data base for planning.

A by-product of the evaluation projects is the identification of certain system features that are not present in UMS but should be provided in operational systems. These are listed as follows:

**operational
systems**

- A facility to interrupt an operation and return control to the user, such as to stop lengthy operations begun in error.
- A continuous indication at the terminal that the system is working, a feature that is useful during lengthy system operations.
- A facility to communicate with the host environment to send and receive messages to/from the central operator.
- A facility to recall earlier input and/or output (automatically available if the terminal is a hard-copy device like a typewriter, but not if it is a visual display unit).
- A facility to output from the system to a data set, thereby making data available to another application.
- A facility to allow multi-user access to the same data base.

Some of these facilities would be simple to add to UMS with varying degrees of effort and little or no redesign of the system. In a research environment these are deficiencies to be lived with and learned from. It was not a function of the UMS project to investigate general interactive system problems.

A facility to communicate with the host environment is outside the scope of this project since it depends entirely upon the operating system environment. Some operating systems (e.g., VM/370) provide this capability.

Facilities to interrupt an operation and to allow multi-user access are the most significant from the system design viewpoint, and are not peculiar to UMS. To provide multi-user access, a comprehensive data base capability is required in which UMS could operate in the same way as it does now. The interruption facility does not involve such a radical change. Since, however, the data access operations are at the lowest level in the system (i.e., farthest from the user), the interruption capability must be inherent to these routines before the feature can be provided in the total system.

from
research
to
product

The results of the evaluation process, as just described, have encouraged the transfer of the research and experience into the development of a Country Implemented Program. It is now a product of IBM United Kingdom Limited that is known as the Interactive Planning System Extended (IPSX/370).⁹ This is described as an interactive data analysis tool rather than as a data base system. Emphasis is placed on its ability to provide users with an unstructured problem solving ability based on copies of operational files that can be updated as the need arises. This need not be too frequent, since, owing to the strategic nature of the work, it may not matter whether the data are one or two weeks (say) out of date.

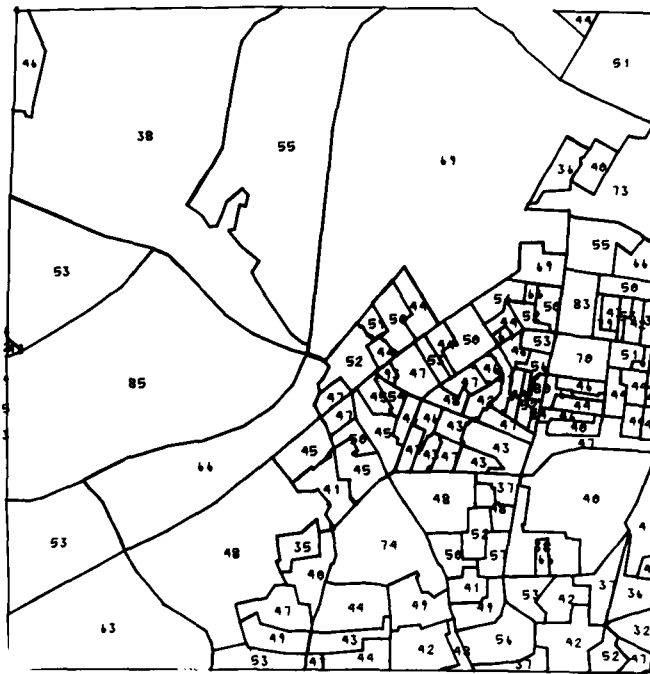
One key feature of IPSX is its ability to give users access to data without the heavy overhead in time and resources involved in writing programs. One user, for example, had a set of data that had been collected two years previously but not processed owing to the unavailability of programming and computer resources. Within a few days, using IPSX, the data had been loaded and answers to his queries obtained. Parts of IPSX have been issued independently as the Spatial Processing Routines.¹⁰ This is a set of input/output-free subroutines that are designed to be called from user-written programs to perform geographic data processing (e.g., point-in-polygon determination and intersection of two areas).

The language extensibility feature of UMS has not been provided in the transferred product because it was believed that the average user would not take advantage of this facility, and therefore the implementation cost was not justified.

Further research

In mid-1976, a further research project concerned with the overall needs of physical planning was begun at the United Kingdom Sci-

Figure 4 Enumeration districts showing percentage of population over retirement age (1971)



entific Centre. Whereas the UMS project considers data analysis, the physical planning project (as it is called) is designed to consider all aspects of planning systems.

By taking the IPSX system as a basis on which to build, the needs of the other stages of the planning cycle are being addressed, including the following:

- Presentation of results, especially graphically
- Incorporation of models
- Provision of statistical routines

For graphic output, an IBM 2250 Vector Graphics Display is being used to display maps and the IBM program product GRAPHAGE has been incorporated to provide histograms, graphs, and symbolic maps on the central or local printer and on the IBM 3277 Display Unit. The next extension envisaged is the use of color and programmable character sets to provide better quality of information display.

graphic
presentation

Figure 4 has been prepared from a photograph of a display on an IBM 2250 display screen. In this instance of spatial distribution of a variable, the variable represents the percentage of the population in an area who have retired from work. These data have been obtained from the 1971 United Kingdom Census of Popu-

lation. The areas whose boundaries are displayed represent enumeration districts, which are the basic units of the Census. Although the map has no identifying features (a fact that makes it unsuitable for the casual observer), it presents no problem to the professional, who can identify the elements of the display because of his familiarity with the area. As long as the display is required only as a working medium and not for final presentation, no improvement of the display needs to be programmed other than windowing and scrolling. Other variables that relate to any area can also be displayed, preferably on a separate alphanumeric screen (e.g., IBM 3277) by pen-detecting the border of the desired area.

models

Mention was made earlier in this paper of the complex interrelationships among the various subject matters involved in planning. Analysis of such relationships can either be left to the educated thought processes of the experienced planning professional, or they can be explored by the creation of mathematical models. Two of the most common models used in planning are the population forecasting model, based upon the cohort survival method, and the allocation or "gravity" models, based upon the theory devised by Lowry, and after whom the models are named.¹¹ A typical *Lowry model* allocates population to a network of zones on the basis of given criteria, notably basic employment in each zone and the degree of attraction between each pair of zones. This attraction depends upon a number of factors, especially journey times between zones and housing facilities in each zone.

If such models are developed, they must be validated or calibrated against known conditions. The development of such models in a planning context is not trivial. One county in the United Kingdom spent ten man-years in developing a residential allocation model, and this did not include its calibration. Sixty percent of the effort was spent in creating the necessary data base to serve as input to the model.

Once the existing framework has been analyzed, and the various interrelationships formulated, the next stage is to forecast future trends based upon a variety of assumptions that are implicit in the various policies being proposed. Again, these predictions can vary between those based on human intuition and those produced by the use of models.

The two main drawbacks to the use of models in physical planning are a lack of appreciation of the models' logic, and their data requirements. Models are generally developed by academics or research groups, probably in conjunction with a planning authority. The continued use of the models, however, is left to the planners. Consequently, unless the assumptions built into a model are

very clearly defined, the use of models to explore future situations is suspect because the necessary requirements are not always met. This is aggravated if such a model is run in batch mode, as is usually the case, with the results received in isolation and often badly presented.

The second problem occurs in the provision of input data for the model. Models are usually developed as stand-alone routines whose input requirements are specified without thorough consideration as to how the data exist. Thus the manipulation of data as input to, output from, or between models can present severe problems to the potential user.¹²

The emphasis in current research at the United Kingdom Scientific Centre is to determine how models can fit into the data analysis system. We seek to rely upon the system to provide the data for input to the model and to hold the resultant output. To this end, each model requires a preprocessor and a postprocessor for relief from input-output considerations. These processors would also prompt the user dynamically for descriptions of the inputs to be used and what to call the results. We believe that the interactive use of models within a data analysis system that gives the facility to explore both the input and the output will greatly improve confidence in the models themselves. By the analysis of repeated operations, and the use of graphic output, interactive use of models should lead to a greater understanding of each model's logic.

The third extension is to give the system a statistical capability, lack of which was one of the weaknesses of the earlier data analysis system. To provide improved statistical capability, the IBM Scientific Subroutine Package (SSP) has been incorporated. This provides many routines (e.g., calculation of means, standard deviations, correlation, and regression analysis), and, like the models, the data input and output are handled by the data analysis system. Thus the data analysis facilities have been increased, and the models can be calibrated.

statistical
capability

To provide the data analysis (IPSX), graphics (GRAPHAGE), statistical capability (SSP), and model routines all in one system module would be unwieldy and inflexible. It would be difficult, for example, to develop and add a new model. Therefore, using VM/CMS facilities, a dynamic fetch capability has been added to IPSX that enables the extra functions (graphics, statistics, and models) to be held in program libraries (CMS TXTLIBS or TEXT files). An extra function is then loaded only as and when it is required. Thus the new prototype looks very similar to that shown in Figure 2, which illustrates the original data analysis system. The only addition is that the FUNCTIONS section now has a dynamic link to a program library on disk.

This implementation has been evaluated internally by another planner working at the United Kingdom Scientific Centre on a fellowship.¹³ Our intention is to extend our research into the remaining stages of planning—the evaluation of alternative policies and monitoring—and then to evaluate the whole system in a joint project with a planning authority.

Relation to other work

The introduction to this paper briefly mentions that the emphasis of the United Kingdom Scientific Centre has been on advancing the application of computers to physical planning. Elsewhere in IBM, there are other groups active in similar areas. Closely related systems are the following:

- Computer Assisted Regional Planning System (CARPS), which aims toward the integrated operation of data, models, and applications in the context of regional planning.¹⁴ This work is being done at the IBM Japan Scientific Centre.
- Generalized Geographic Information System (GGIS), the objectives of which are compatibility with existing geographic base files, such as DIME files, applicability to demographic and resource data, application independence, and geographic data structure compatibility.¹⁵ This was a project of the IBM Federal Systems Division.
- Geo-data Analysis and Display System (GADS) emphasizes the solution by nonprogrammers of problems involving data that can be related to a geographic location. GADS provides a data extraction technique for accessing data in a variety of files, and a set of conversational data analysis and display functions.¹⁶ This work is being carried out at the IBM Research Laboratory at San Jose, California.

It is instructive to note that CARPS, GADS, and UMS/IPS all adopt the relational, or tabular, approach to data management, whereas GGIS is based upon the hierarchical approach, using DL/I.¹⁷ Thus it would seem that both approaches are feasible, although the interactive systems favor the relational methodology.

Performance problems associated with large amounts of data were tackled in GADS by using data extraction and aggregation techniques, whereas the experience with UMS/IPS is that this problem is a necessary overhead in order to give the user complete freedom to access the data and to aggregate it in whatever way he chooses. Each of these projects can benefit from the experience of the other. The variety of approaches reflects the variety of experience and requirements of the regions in which the studies are being carried out.

Specialized studies have been undertaken, in the Research Division particularly, that have concentrated fully upon specific sub-systems such as the following: graphics,¹⁸ query language,¹⁹ and decision support systems.²⁰ Each of these research areas has advanced the state of the art to a great degree. UMS/IPS has added to this fund of knowledge by seeking a deeper understanding of needs in the application area. The long-term potential for physical planning systems is positive. We have concluded that among the immediate considerations of hardware and software cost should also be included a simplified initial approach to physical planning.

Concluding remarks

This paper has discussed the development of a geographic physical planning system and its evaluation by users in two modes of application—one that emphasizes problem analysis and solution and one that emphasizes querying a practical data base. Results have been sufficiently encouraging that the experimental system has formed the basis of an operational system product. The initial results have also been sufficiently successful that further research is being pursued toward the addition of improved graphic presentations, city planning models, and statistical capabilities. Basic research and operational systems are being studied elsewhere in the world with good results. Each system, including the Urban Management System presented in this paper, adds to the general fund of knowledge from which all can draw. Each system also aims toward meeting the special needs of the home region in which it was developed.

CITED REFERENCES

1. B. K. Aldred and B. S. Smedley, *An Urban Management System—General Overview*, IBM United Kingdom Scientific Centre Report /53, The Information Officer, IBM United Kingdom Scientific Centre, Neville Road, Peterlee, Co. Durham, England (May 1974). (ITIRC AAA 74A003331.)
2. B. K. Aldred and B. S. Smedley, *UMS Technical Overview*, IBM United Kingdom Scientific Centre Report /50, The Information Officer, IBM United Kingdom Scientific Centre, Neville Road, Peterlee, Co. Durham, England (February 1974). (ITIRC AAA 74A002800.)
3. United Kingdom Department of the Environment, *GISP—General Information System for Planning*, Her Majesty's Stationery Office, London (1972).
4. United Kingdom Department of the Environment, *Manual of Point Referencing*, Her Majesty's Stationery Office, London (1972).
5. S. H. Mandil, *MP/3 Macro Processor*, IBM United Kingdom Scientific Centre Report /44, The Information Officer, IBM United Kingdom Scientific Centre, Neville Road, Peterlee, Co. Durham, England (December 1973). (ITIRC AAA 74A003024.)
6. S. J. P. Todd, "The Peterlee Relational Test Vehicle—a system overview," *IBM Systems Journal* 15, No. 4, 285–308 (1976).
7. B. S. Smedley, *UMS and Geographic Data Processing in Urban Planning*, IBM United Kingdom Scientific Centre Report /79, The Information Officer, IBM United Kingdom Scientific Centre, Neville Road, Peterlee, Co. Durham, England (April 1976). (ITIRC AAA 76A002335.)

8. J. S. M. Verhofstad, *The PRTV Optimizer—the Current State*, IBM United Kingdom Scientific Centre Report /83, The Information Office, IBM United Kingdom Scientific Centre, Neville Road, Peterlee, Co. Durham, England (May 1976). (ITIRC AAA 77A002017.)
9. IBM United Kingdom Ltd., *Interactive Planning System Extended—Program Documentation and Operations Manual*, Order Number SB10-6516, IBM United Kingdom Ltd., Technical Information Centre, 17 Addiscombe Road, Croydon, Surrey, England.
10. IBM United Kingdom Ltd., *Spatial Processing Routines—Program Documentation and Operations Manual*, Order Number SB10-6514, IBM United Kingdom Ltd., Technical Information Centre, 17 Addiscombe Road, Croydon, Surrey, England.
11. Ira S. Lowry, "Seven models of urban development: a structural comparison," *Urban Development of Models*, edited by George C. Hemmens, Highway Research Board, National Academy of Sciences, Washington, DC (1968), pp. 121–163.
12. A. Gillard and S. Openshaw, "SPL—a proposal for a modelling and data-base system for urban and regional modellers and structure planners," *Planning Outlook* (Planning Press, Oxford, England) **18**, 1–11 (1976).
13. P. B. Hayton, *Physical Planning Project—Evaluation Report*, IBM United Kingdom Scientific Centre Report, The Information Officer, IBM United Kingdom Scientific Centre, Neville Road, Peterlee, Co. Durham, England (to be published).
14. *Concept of Regional Development Planning System*, IBM Japan, Tokyo Scientific Centre Report GE18-1826-0.
15. M. H. Blumberg, *A Generalized Geographic Information System for managing a non-redundant demographic and resource data base*, paper presented at the Urban and Regional Information Systems Association (URISA) Conference, Seattle (1975).
16. E. D. Carlson, J. L. Bennett, G. M. Giddings, and P. E. Mantey, "The design and evaluation of an interactive Geo-data Analysis and Display System," *IFIP Conference Proceedings, Information Processing 74*, North-Holland Publishing Company, Amsterdam (1974), pp. 1057–1061.
17. *Information Management System (IMS), General Information Manual*, Order Number GH20-1260, IBM Corporation, Data Processing Division, White Plains, New York 10604.
18. R. Williams, *A Picture-Building System*, Research Report RJ 1560, IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598. (ITIRC AAA 75A002785.)
19. M. M. Zloof, "Query-by-Example: a data base language," *IBM Systems Journal* **16**, No. 4, 324–343 (1977).
20. B. F. Grace, *Training Users of a Decision Support System*, Research Report RJ 1790, IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598. (ITIRC AAA 76A003104.)

Reprint Order No. G321-5080.