1-1-1988

Design of an expert system for spatial planning

Panagiotis Lolonis

Marc P Armstrong
University of Iowa

DESIGN OF AN EXPERT SYSTEM FOR SPATIAL PLANNING

Panagiotis Lolonis
and
Marc P. Armstrong

Department of Geography
316 Jessup Hall
The University of Iowa
Iowa City, IA 52242

ABSTRACT

Spatial data processing systems are now making increasing use of optimization techniques to help decision-makers locate facilities. Because of several shortcomings of these techniques, we develop an expert system to augment current decision-making processes. The developed system consists of a spatial database, with an accompanying rule base derived from empirical knowledge about facility location. Prolog is used as the inference engine.

INTRODUCTION

The integration of spatial analysis and spatial data handling technologies is rapidly becoming commonplace (Armstrong, Densham and Rushton, 1986; Diamond and Wright, 1988; Lupien, Moreland, and Dangermond, 1987). It is important to note, however, that the spatial analysis models which are the essential components in these decision support applications, have been criticized because they are inflexible, and fail to account for all aspects of complex locational problems. Because of these shortcomings, we have shifted our attention to alternative procedures for solving locational problems, and we have adopted artificial intelligence principles to construct an expert system designed to provide guidance in facility location.

Expert systems provide a useful framework for integrating existing locational problem solving methods. They exhibit numerous useful characteristics, including flexible representation of knowledge, and the capability to handle uncertainty. Unfortunately, disadvantages to the application of expert system procedures exist. Combinatorial explosions sometimes occur as alternatives are explored; also, numerical operations and therefore, optimization, are difficult to perform. For these reasons we employ a staged approach to system development, using a logic programming module for eliminating unsuitable locations, an analytical module for performing site selection, and additional modules to provide assistance in the final evaluation of alternatives. In the remainder of the paper we first describe the problem domain, and then discuss the advantages and disadvantages of purely analytical approaches to locational problems; finally we describe the characteristics of our expert system.

PROBLEM DOMAIN

The problem domain of the system is the location of retail facilities and services. The main components of the problem are areal units which must be evaluated not only in terms of their physical, economic, and legal attributes, but in terms of their relative location as well. The problem is complex because it is combinatorial, and requires the collection, organization, and processing of large amounts of information. To cope with this complexity, decision-makers often perform location selection using a three-stage hierarchical approach (Ghosh and McLafferty, 1987). At the first, and most general scale, experts identify a market area (e.g. SMSA) with a high market potential for the facility being examined. Then, at the intermediate level, experts refine the search, and choose areas (e.g. zip codes) with a high likelihood for demand within the market area selected in the first stage. Finally, within the area(s) selected in the second stage, specific sites are evaluated for facility suitability. At each stage a strong component of subjectivity often is present, although recent advances in analytical methods have enabled the adoption of more objective mathematical models.

ANALYTICAL METHODS FOR SOLVING LOCATIONAL PROBLEMS

Progress in operations research during the past two decades has broadened the array of techniques used for locational problem solving, from locating multiple facilities on a plane or network (Cooper, 1967; Maranzana, 1969), to incorporating spatial interaction (Beaumont, 1980) and consumer welfare functions (Hodgson, 1981). Other progress has been made in including multiple, competing objectives in a modelling framework (Schilling, 1980; Cohon, et al., 1980). Analytical methods have been widely adopted because:

1) Information is provided about important components of locational problems including, the number and location of facilities, and the allocation of demand.

2) The procedures are objective, systematic, and except for only the largest of problems, solution times are reasonable, especially if the problem can be restricted in some way.

3) A framework is provided in which the performance of a facility can be assessed with respect to the remaining facilities.

4) In most cases, required data are easily obtained.

Despite these rather important advantages, negative attributes of analytical models can be discerned:

1) An assumption underlying the use of the models is that decision-makers are able to express their preferences to analysts. Error can be introduced because analysts may distort the problem, both in the process of preference elicitation, and also while casting the problem in mathematical form.

2) A second assumption is that all important aspects of a problem can be captured in a mathematical formulation. Even complex objective functions, however, fail to consider factors which are important in locating facilities, and qualitative criteria are especially difficult to include.

3) Modelling practices invariably introduce error into solutions. These errors are well known, but pervasive (Current and Schilling, 1987).

These factors have persuaded us that efforts must be marshalled to improve the process of facility location within a GIS framework. Artificial intelligence principles can be used to offset these problems, and to augment the capabilities of analytical models.

EXPERT SYSTEM DEVELOPMENT

Expert systems form one branch of the rapidly growing field of artificial intelligence. Although expert systems are not applicable to all problems, our application is appropriate because the domain of the application can be delineated, and the elements, and relationships among elements, can be isolated from those which are irrelevant. The problem is further suited to expert system development because it is decomposable; the evaluation of each candidate site can be broken into smaller problems which can be manipulated separately. For example, to determine whether a site is suitable for a facility, it must be accessible, and must possess appropriate infrastructure support. Access can be decomposed further to examine such factors as: access to public transport, and traffic bearing capacity of roads.

The location of multiple facilities, however, cannot be decomposed easily into simpler independent problems, because the performance of a facility also depends upon the unknown locations of other facilities. This class of problems can be solved if the total facility configuration is evaluated by the expert system. This assumes that the number of alternative configurations (layouts) is relatively small, and that each configuration is identifiable. If, however, the number of alternatives is large, then location-allocation software can be used to reduce the number of alternatives in a preprocessing operation. In that case, the procedural software is used to winnow unacceptable sites on the basis of the objective function values.

The problem domain of facility location has accumulated a substantial amount of empirical knowledge (e.g. Ghosh and McLafferty, 1987), which can be used to structure rules to guide system development. A portion of this knowledge has led us to conclude that mathematical formulations alone will not lead to the solution of all locational problems because relationships among factors are not always certain, and qualitative criteria are used by decision-makers.

Approach to Problem Solution

A general purpose system is difficult to specify because different types of facilities use different rules, and also, factors are weighted differently for each type of facility at each scale of analysis. At the beginning of each analysis the user is asked about the type of facility to be located, and the scale at which the analysis will be performed. Following this, a three stage process is undertaken:

1) Evaluate selected spatial units with respect to their suitability as facility sites. Evaluation may be expressed along one or more dimensions; more than one dimension will always be used when multiple objectives are examined. The level of measurement can be binary (acceptable, unacceptable), symbolic (good, fair, poor), ordinal, or interval scaled.

2) If there are many highly rated sites, and all possible configurations cannot be evaluated, then use optimization techniques, with an appropriate objective function, to reduce the number of alternatives. Optimization, however, may be required only for multiple location problems; for single location problems, the results from the first stage should be sufficient to provide a recommendation.

3) Evaluation of near optimal configurations is often performed in multi-attribute location selection. AI methods are used to evaluate facility layouts which are optimal (or near optimal) solutions from the second stage.

System Components

The system is implemented in two related programming environments- PROLOG and Pascal. The combination that we are using allows us to activate analysis modules written in Pascal from the logic programming environment. In this way, numerical processing operations can take place outside of the logic program, although the result of each operation can be applied in the declarative domain. The system consists of a user interface, an inference engine, knowledge base, procedural processing module, graphics, and tabular report generator (Figure 1). Each component is described below.

User Interface. The initial interface is menu-based, and is written in PROLOG. User input is requested only when the proper option is selected, and responses are echoed to the user to verify input. A natural language interface was rejected temporarily, because of the expense of developing the software, and the increased potential for input error. Because of their great flexibility, however, a natural language interface may be added at a later time.

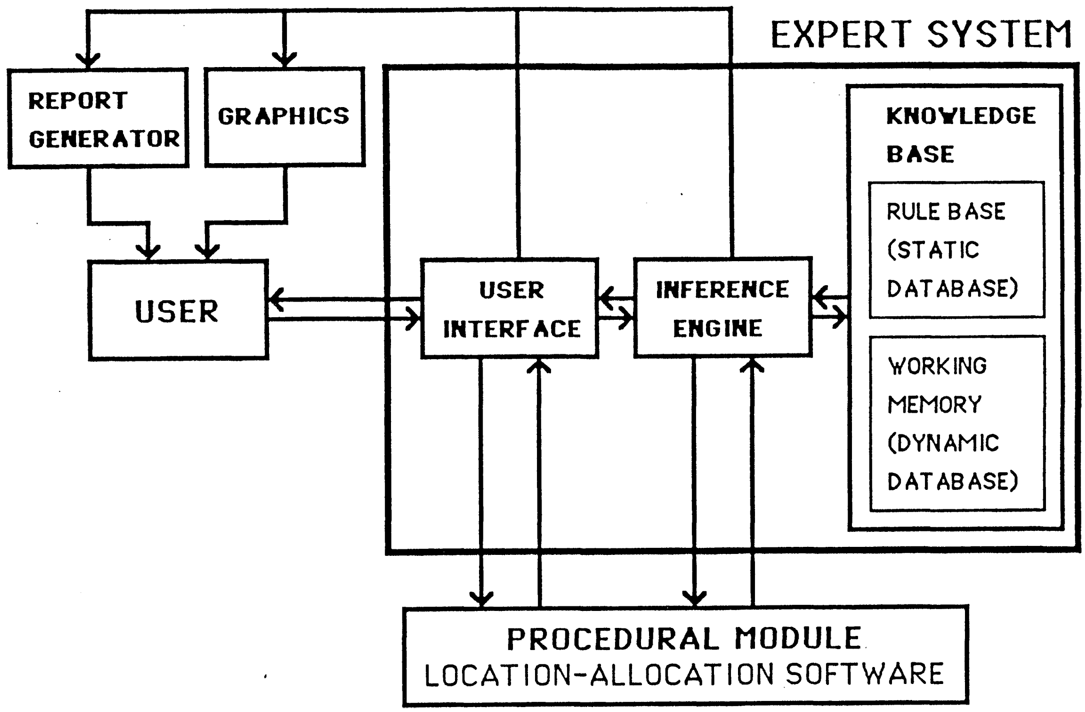


Figure 1. System components.

<u>Inference Engine</u>. The inference engine performs two tasks: controlling the order in which rules are executed, and drawing inferences from rules. PROLOG is used as the inference engine because:

1) It supports backward chaining for directed searching, which reduces the likelihood of combinatorially explosive searches. Additionally, it supports backtracking when rules fail to succeed.

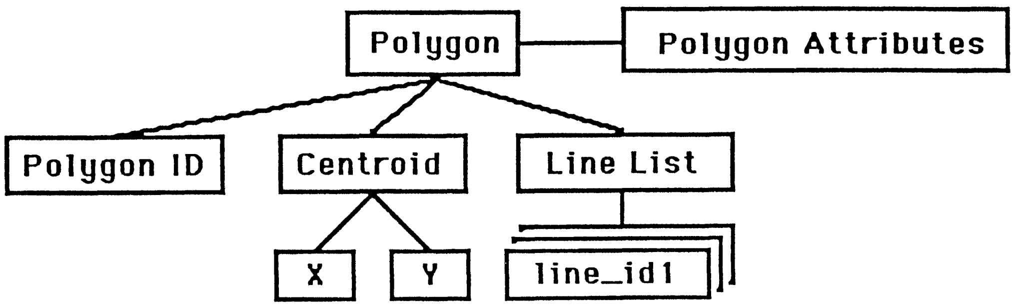
2) It is based on the principle of **resolution refutation**, which is well established in mathematical logic.

3) It is less error prone in adding new rules than other inference engines such as EMYCIN, and it facilitates the development of flexible rules (Jackson, 1986: 175).

In the system, the inference engine is functionally connected with the knowledge base from which it retrieves information; knowledge is processed and new facts are stored in working memory. The inference engine also directs output to a report generator, and to a graphics module to help decision makers evaluate solutions, or processes that lead to solutions. Finally, the inference engine activates analytical procedures.

<u>Knowledge Base</u>. The knowledge base contains all available facts and rules which are used in the inference process. The knowledge base has two components: the rule base and working memory.

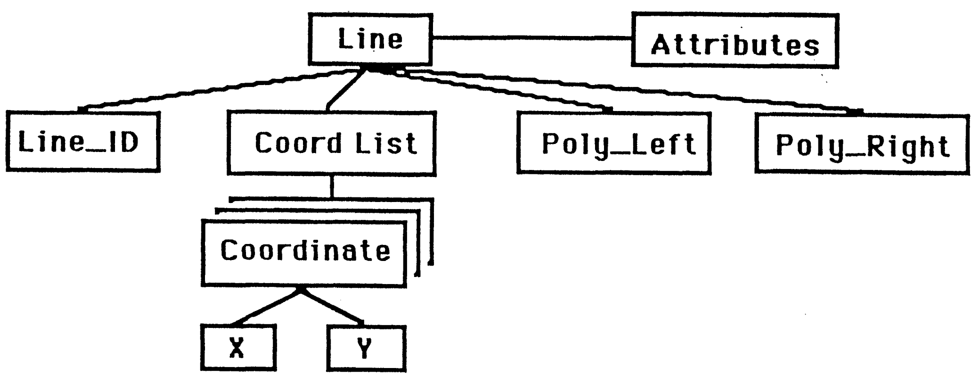
Rule base. The rule base contains static **facts** and **rules** which are unaltered in the consultation process. In the facility location domain, facts consist of knowledge about entities defined geometrically as chains and polygons. Depending upon the scale of analysis, areal entities may be SMSAs, census tracts, or individual parcels. Topological and attribute information is also associated with each object (Figure 2). The first argument of the polygon structure is a polygon ID. The second argument is a list containing coordinates of the polygon centroid which is used for proximity analyses. The third polygon argument is a list of chain identifiers that describe its perimeter for map displays. Finally, attribute data consist of such items as physical, and economic characteristics of each entity which are used to assess suitability for facility locations.



Polygon (Polygon ID, [x_coord, y_coord], [line_id1,... line_idn], Attributes).

Figure 2. Logical data structure for polygon entities.

A second level of structure is needed to build chains from coordinate lists (Figure 3). In addition to geometrical information, we also include topological and attribute information (e.g. border, road) for each chain. Topological relationships are used to assess neighborhood polygon status, especially at the site scale of analysis.



Line (Line_ID, [[X1,Y1],...[Xn,Yn]],Poly_LeftID,Poly_RightID, Attributes).

Figure 3. Logical data structure for linear features.

The second component of the rule base is the set of rules from which inferences are drawn. Each rule consists of a number of conjunctive (or disjunctive) conditions, and a conclusion. Conditions are predicates which may be found as facts in the knowledge base, or their veracity can be determined from user input. The conclusion of a rule is a PROLOG premise which is asserted in working memory, if all the conditions of that rule are proved to be true during the inference process.

An example of our rule structure is shown in Figure 4. The first rule states that a polygon receives an excellent rating for locating a facility, if it has excellent physical and economic attributes. The second rule specifies that a polygon has excellent physical attributes, if its size ranges between 50,000 and 90,000 square feet, and is vacant. The third rule states that a polygon has excellent economic characteristics, if its land value is less than $5000 and the prevailing tax rate is less than 7%. The final two rules describe the certainty with which the conclusion of the first rule is asserted in the knowledge base.

```
/*  EXAMPLE RULE FOR RATING POLYGONS  */
  site_rating( Polygon_ID, excellent, CF ) :-
    polygon( Polygon_ID,_,_,_,_,_,_ ),
    physical_attributes( Polygon_ID, excellent, CF1 ),
    economic_attributes( Polygon_ID, excellent, CF2 ),
    compute_certainty( [CF1, CF2], 800, CF ). /*CF=800 */

/*  RATING POLYGONS ON PHYSICAL ATTRIBUTES */
  physical_attributes( Polygon_ID, excellent, 1000) :-
    polygon( Polygon_ID,_,_,Size, Situation,_,_),
    Size > 50000, Size < 90000, Situation=vacant.

/* RATING POLYGONS ON ECONOMIC ATTRIBUTES  */
  economic_attributes( Polygon_ID, excellent, 900 ) :-
    polygon( Polygon_ID,_,_,_,_,Land_Value, Tax_rate ),
    Land_Value < 1000, Tax_rate < 7.

/*  RULES FOR COMPUTING CERTAINTY FACTORS  */
  compute_certainty( [],_,_ ).
  compute_certainty( CF_List, Rule_CF, Conclusion_CF ) :-
     min_element( CF_List, Min_element ),
     Conclusion_CF = Rule_CF * ( Min_element / 1000 ).
```

Figure 4. Code fragment to illustrate rule structure.

Working memory. Working memory forms the dynamic component of the knowledge base. At the start of each consultation, working memory is empty, but new facts are added to it as rules are fired, and inferences are made. In addition to inferred facts, user responses are also stored in working memory. This speeds interaction because the system does not ask questions that have already been answered. The facts in working memory have the same structure as facts in other system components, and reside in working memory for the duration of the consulting session.

Procedural Module. This module includes procedures written in Pascal which are activated from the user interface in the PROLOG environment. They perform tasks requiring large amounts of numerical calculations (e.g. objective function optimization in location-allocation). The procedural module interacts with the inference engine and user interface for accessing information.

Report Generator and Graphics Module. The recommendations of the expert system cannot be fully evaluated unless they are explained in detail. The large numbers of rules used in each consultation results in lengthy explanations of the reasoning process, and that makes essential the generation of easy to interpret reports. The report generator, therefore, is used to format consultation results for decision-makers.

Graphic displays are used to plot maps and graphs to improve human-computer interaction. The user not only has the capability to view information, but also to provide information. For example, the user may choose polygons to be evaluated as possible facility locations, or may request information about a particular polygon or line.

DEALING WITH UNCERTAINTY

In locational problems, there are several sources of uncertainty such as measurement error, and dynamic changes to actions or plans of competitors. In this section, however, we focus on uncertainty introduced by the formulation and interaction among rules. Although many expert systems have mechanisms to deal with uncertainty, all aspects of the issue have yet to be resolved. Probability theory, which is possibly the most rigorous and well founded frame for dealing with uncertainty, is not normally implemented in AI applications, because assignment of probabilities involves the use of mathematical techniques which are not implemented easily in declarative programming languages.

Because rules are based on imperfect human expertise, they are not expected to always hold. Thus, even when all conditions of a rule are true, there might be more than one possible incompatible conclusion that can be inferred. Also, the arguments of some rules may be inexact, or fuzzy; inexactness arises from qualitative evaluations, and as a result, conclusions must be assigned an appropriate level of uncertainty. Specifically, certainty of the conclusion of a rule is measured by a **certainty factor**, which is an integer ranging between 0 and 1000. It must be noted that the range is somewhat arbitrary; other intervals such as [0,1] could be used without undue effect on the inference process. In our system, uncertainty is treated in a way similar to that adopted in MYCIN (Jackson, 1986). For example, given that all conditions of a rule are true with some level of certainty, and that the conclusion has a certainty factor of zero, then there is total uncertainty about the truth of a rule, and we assert that it has failed. On the other hand, if all conditions of a rule are true and the certainty factor associated with the rule is 1000, then the conclusion is certainly true.

The computation of uncertainty is based upon the following definitions. If the premise of a rule has certainty C(1) and the certainty factor associated with that rule is C(2), then the certainty factor of the conclusion of the rule is calculated as: $C = C(1) * (C(2)/1000)$. If, for example, the rule has the general form:

IF condition THEN conclusion CF = 800

where CF is the certainty factor, and there is evidence in working memory that the certainty of the condition is C(1)= 900, or 90% certain, then the conclusion is asserted in working memory with certainty factor:

$$C = 900 * (800/1000) = 720.$$

In this case, there is a 72% certainty that the conclusion is true.

The condition part of the rule can be either a single condition, or a composite condition formed from several single conditions joined with conjunctions and disjunctions. When a composite condition formed from conjunctive premises is encountered, then the certainty factor of the whole condition is set equal to the lowest certainty factor of the premises. When premises are joined with disjunctions, on the other hand, then the composite certainty factor is equal to the largest certainty factor of the premises. Facts drawn from the rule base are absolutely certain (C = 1000), as are answers supplied by system users, unless otherwise specified.

CONCLUSIONS

By integrating GIS, operations research, and artificial intelligence methods, our approach shows promise for overcoming the limitations of methods currently employed in locational planning. Each technology lends assistance in overcoming different aspects of locational problem solving. Operations research methods reduce the solution space for locational problems by providing optimal, or near optimal, locational configurations to decision-makers. GIS provides the means for organizing and displaying spatial information to decision-makers. Finally, artificial intelligence methods enable the representation of accumulated empirical knowledge about facility location problems, and permits the manipulation of such knowledge contained in a knowledge base. This combination of components allows alternative recommendations to be compared, and provides human decision-makers with a powerful set of tools for enhancing their locational problem solving capabilities.

REFERENCES

Armstrong, M.P., Densham, P.J., and Rushton, G. 1986. Architecture for a microcomputer based spatial decision support system. Proceedings, Second International Symposium on Spatial Data Handling. Williamsville, NY : IGU Commission on Geographical Data Sensing and Processing.

Beaumont, J.R. 1980. Spatial interaction models and the location-allocation problem. Journal of Regional Science, Vol. 20, No 1, 37-50.

Cohon, J.L., ReVelle, C.S., Current, J., Eagles, T., Eberhart, R., and Church, R. 1980. Application of a multiobjective facility location model to power plant siting in a six-state region of the US. Computers and Operations Research, Vol. 7, Nos 1 and 2.

Cooper, L. 1967. Solutions of generalized locational equilibrium models. Journal of Regional Science, 7, 1.

Current, J.R., and Schilling, D.A. 1987. Elimination of A and B errors in p-median location problems. Geographical Analysis, Vol. 19, No 2.

Diamond, J.T., and Wright, J.R. 1988. Design of an integrated spatial information system for multiobjective land use planning. Environment and Planning B: Planning and Design, Vol. 15, pp 205-214.

Ghosh, A., and McLafferty, S. 1987. Location Strategies for Retail and Service Firms. Lexington, MA: Lexington Books.

Hodgson, M.J. 1981. A location-allocation model maximizing consumers' welfare. Regional Studies, Vol. 15, No 6, pp. 493-506.

Jackson, 1986. Introduction to Expert Systems. Wokingham, England: Addison-Wesley.

Lupien, A.E., Moreland, W.H., and Dangermond, J. 1987. Network analysis in geographic information systems. Photogrammetric Engineering and Remote Sensing, Vol. 53, no. 10, pp. 1417-1421.

Maranzana, F. 1964. On the location of supply points to minimize transport costs. Operations Research Quarterly, Vol. 15, No. 3, 261.

Schilling, D.A. 1980. Dynamic location modelling for public sector facilities: a multicriteria approach. Decision Science, pp. 714-723.