Development and Implementation of a Knowledge-Based GIS Geological Engineering Map Production System

*Jill J. Cress**
Department of Computer Science, University of Missouri-Rolla, Rolla, MO 65401
Robin R.P. Deister
Department of Geological and Petroleum Engineering, University of Missouri-Rolla, Rolla, MO 65401

ABSTRACT: A research project to automate Geological Engineering Map (GEM) production resulted in the development of a knowledge-based GIS (KBGIS) using the GoldWorks expert system shell. Implementation of the rule base produced a valid GEM, but required a significant amount of production time. An effort to find a more efficient implementation resulted in the creation of a Conversion Expert System (CES). The CES accepts a GoldWorks KBGIS as input, and produces the equivalent FORTRAN code as output. This FORTRAN can then be used to produce accurate GEMs in significantly reduced production time. Therefore, a two-step process has been created for the development of GEM production systems. First, use GoldWorks for the rule base development, and second, use the CES to convert the KBGIS into FORTRAN for implementation. This two-step process provides an easily developed and maintained knowledge base, as well as an efficient GEM production system utilizing a GIS approach.

INTRODUCTION

EXPERT SYSTEMS are rapidly becoming an integral part of almost every field. This trend can be traced to the growing need to add "knowledge" to processes and the proven success of using expert systems to fulfill this need. Recently, in fact, the need for a more knowledgeable and efficient production process led to the introduction of expert systems into the area of Geological Engineering Map (GEM) production.

GEMs represent the "geological engineering conditions" of an area; they "show surficial and bedrock geologic patterns classified according to engineering suitability for urban development. . . ." (Usery *et al.*, 1988a; Usery *et al.*, 1988b). GEMs should portray objective information in order to best evaluate the engineering involved in regional planning. The specific information on any single GEM is represented by the chosen classification scheme, which, in turn, is based on the intended use of the map. One of the primary uses of GEMs is as a tool in determining a suitable site for a particular type of development. For example, if a possible landfill site were being investigated, geological engineering characteristics such as flooding frequency and soil plasticity index would be two important parameters. Therefore, the GEM produced would classify the area under investigation in terms of these parameters.

A classification scheme refers to the structure used to organize the Earth resource parameters of a GEM. Each classification scheme is developed by determining what parameters are to be represented on the map, and then defining each classification unit in terms of the chosen parameters. Parameters can refer to general properties such as agricultural soil type, bedrock geology, and elevation, as well as derived attributes of these properties, i.e., flodding frequency, permeability, presence of karst, plasticity index, and percent slope.

One important requirement for any GEM is that it be easily understood. Information should be presented in a manner that meets the needs of the end user and portrays the information needed to evaluate the situation. Therefore, after choosing a classification scheme, how the information is to be presented on the finished map must also be determined. One presentation method for a specific site selection map is an "all or nothing" product. This format would identify only those areas meeting all of the favorable criteria, resulting in only two classes; it meets the criteria or it doesn't. Another presentation method would be to classify all areas by showing what parameters were present.

GEM production was automated by utilizing a knowledge-based GIS (KBGIS) approach. This involved the development of a KBGIS and then the creation of another expert system, which is used to convert the KBGIS into FORTRAN code (Figure 1). In other words, a two-step process has been created for the development of efficient automated GEM production systems that utilize GIS data. The first step is to develop the necessary geological engineering knowledge base using the GoldWorks[†] expert system shell, and the second is to convert the knowledge base into FORTRAN code for implementaiton as the production system. These two steps work together to achieve the goals of an easily developed and maintained knowledge base, as well as an efficient final production system which utilizes GIS data.

APPROACH

Development of this system proceeded in the following stages. First, an in-depth study of the current manual GEM production method was performed. This permitted identification of any weaknesses inherent in the system, in addition to providing insight into the best method of automation. Developing an automated system was the second development stage. A KBGIS approach to automation was chosen, and an appropriate KBGIS was developed and implemented in the GoldWorks environment. It was discovered that the implementation method chosen for the KBGIS was not time efficient. Although the KBGIS approach did solve the manual production system weaknesses, a more efficient implementation method was needed. This led to the third, and final, development stage which involved the creation of another expert system using the GoldWorks environment. This new expert system converts a KBGIS, developed in GoldWorks, into FORTRAN code. This FORTRAN code, after it has been compiled and linked, can then be used as the GEM production system, resulting in siginificant production time savings.

*Presently with the U. S. Geological Survey, 1400 Independence Road, Rolla, MO 65401.

[†]Any use of trade names and trademarks in this publication is for descriptive purposes only and does not constitute an endorsement.

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING,
Vol. 56, No. 11, November 1990, pp. 1529–1535.

0099-1112/90/5611–1529$03.00/0

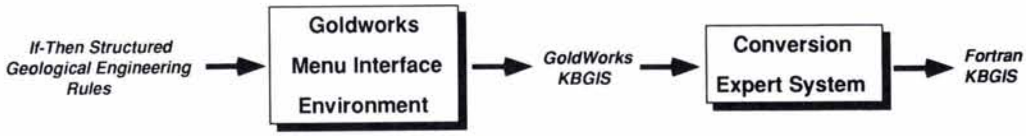


FIG. 1. Two-step process for the development of automated GEM production systems.

STUDY OF CURRENT MANUAL GEM PRODUCTION SYSTEM

GEM production begins with the identification of an area to be mapped. Next, a series of six processing steps are performed; most, if not all, of these steps are done manually. In addition, these steps must be performed each time a GEM is produced, and are usually done by a geological engineer. In fact, the first three steps rely heavily on input and interpretation from an engineer skilled in both geological principles and mapping concepts.

The first step in the current GEM production process is to establish the specific criteria. In the previously cited example of the landfill site selection map, flooding frequency and soil plasticity index are two parameters required for the GEM. The engineer would develop the criteria by stating that the landfill site must not be inundated by flood waters and must not contain plastic soils. In step two, relevant maps of the area are acquired, and in the third production step, the maps are analyzed by the engineer to identify suitable sites. Printing the map is the fourth step in the production process. The engineer is not required at this step because additional analysis is not required during printing. Because the printing process is not automated, however, considerable time can elapse before a usable product is produced. The fifth production step is determining acceptance of the GEM. This involves not only confirming that the map meets map accuracy guidelines, but also reaffirming the criteria. If any problems are identified at this point, or if the users request any changes in the criteria to be used, the production process must be repeated until the map is finally both approved and accepted. Although the map has been accepted, it is not considered complete until it has been field checked, which is the last step in the production process. Field checking involves verifying the new map by going into the field to check it. This check is performed by confirming that the classification for a specific area on the map matches the actual geological conditions. If any discrepancies are discovered, then the necessary production steps must be repeated, including gaining the users' approval, until the map is finally determined to be correct. Once the GEM has been successfully field tested, the production process is complete and the final GEM is produced.

Several weaknesses were observed in the manual GEM production system. One of the more serious is the lack of standardization. Because there are no set policies or guidelines for the development of classification schemes, it is entirely possible to discover two maps which use the same parameters in different ways (Lutzen and Rockaway, 1970; Lutzen and Rockaway, 1971). These classification schemes are obvious victims of the lack of standardization. Subjectivity is introduced even though each engineer applies the same basic geological engineering principles when developing these classification schemes. This subjectivity occurs when the engineer uses his/her own personal knowledge and interpretation techniques to define the classification scheme. The result of this subjectivity is a severe weakening of the power of the different classification schemes. This weakening becomes evident by the lack of standardization in classification nomenclature. For example, one map might define Class Ib as flood prone whereas another map would state that Class Ib lacks flooding.

Another weakness inherent in the manual GEM production process is the demands it places, in terms of both time and effort, on a skilled geological engineer. Few of the steps can be performed without extensive interaction with the engineer, which makes the process extremely labor intensive. The production process also uses an iterative approach for all of its resource gathering and data manipulation. This means that not only does the engineer have to spend a large amount of time developing each GEM, but also that he/she will be performing the same basic processing steps repeatedly until the product is accepted and approved. Because of the demands this process places on a geological engineer, much time is spent on routine activities instead of skilled, creative activities.

In summary, although the current manual GEM production system fulfills its purpose by producing a GEM, it does so in neither an efficient nor reliable manner. In fact, the weaknesses inherent in this system – the lack of standardization and the demands placed on the skilled engineer – make it a prime target for modernization.

DEVELOPMENT OF THE KBGIS

The weaknesses found in the manual GEM production system could be significantly improved with the development of an automated production system. Automation would best be served by utilizing expert knowledge in a GIS environment. This rule-based system has the ability to interact with existing GISs, and is known as a knowledge-based GIS (KBGIS) (Usery *et al.*, 1988b; Smith and Pazner, 1984). A joint research project to develop a KBGIS for GEM production was started in 1987 by the University of Missouri-Rolla (UMR) Department of Geological and Petroleum Engineering and the United States Geological Survey (USGS). The first objective of this project was to identify the basic Earth resource parameters "needed to make engineering judgments about areas... ." (Usery *et al.*, 1988b). Figure 2 provides a generalized flowchart of the procedures used for the development of the KBGIS.

During the initial investigation of exisiting maps, two important facts were discovered. The first was that the GEM classifications varied dramatically from region to region. It was clear that the development of a single set of valid classes would be virtually impossible. As a result of this, the decision was made to narrow the scope of the investigation to include only the Midwestern region. The second discovery was that no standardization existed among the different classification schemes, although there were two basic formats of geological engineering maps. One GEM format provides a general statement of the engineering properties of an area. This general GEM is typified by the Creve Coeur quadrangle GEM that was developed by Lutzen and Rockaway (1970). The other type of GEM format provides specific judgments about the engineering properties in terms of an engineering application. For this project, a general GEM was chosen as the desired end product. The general GEM increases the efficiency of the production system by removing the need to repeat the entire production process for each specific site selection GEM because a one-step translation from a general GEM to a specific GEM can easily be performed.

In an effort to determine the basic Earth resource parameters necessary for the creation of a GEM, a sample of five GEMs was chosen for a detailed investigation. This investigation consisted

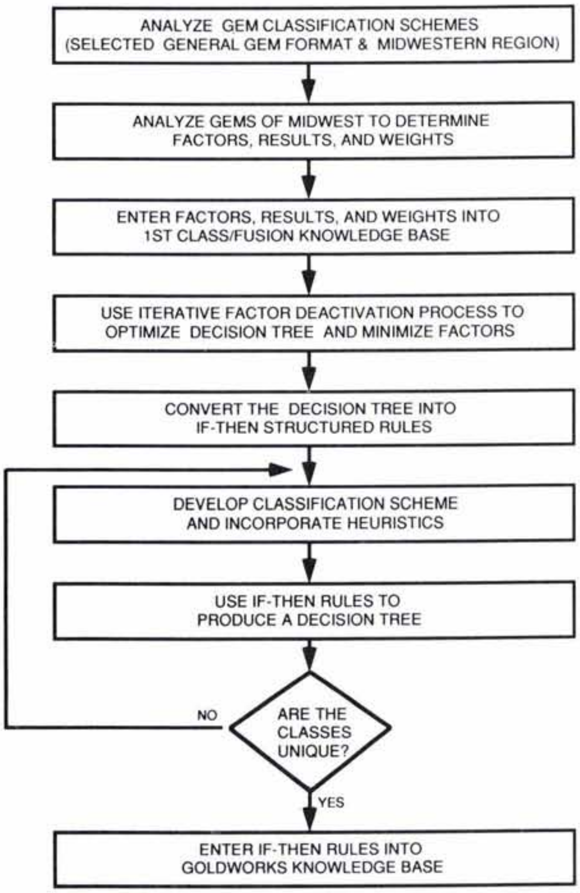


FIG. 2. Procedure for GEM knowledge base creation.

GEOLOGY	FLOODING	KARST	PLASTICITY INDEX	SLOPE	PUBLISHED MAP CLASS	KBGIS CLASS	WEIGHT
ALLUVIAL	YES	NO	20.0	2.	Ia	1	2.00
ALLUVIAL	***YES***	***NO***	***10.0***	***2.***	***Ib***	***2***	***2.00***
ALLUVIAL	OCC	NO	10.0	2.	Ic	4	2.00
ALLUVIAL	NO	NO	20.0	20.	Id	10	2.00
ALLUVIAL	NO	NO	27.0	9.	Ie	5	2.00
LIMESTONE	NO	YES	15.0	25.	IIa	6	2.00
LIMESTONE	NO	YES	32.0	20.	IIb	7	2.00
LIMESTONE	NO	YES	15.0	20.	IIc	6	2.00
LIMESTONE	NO	YES	20.0	40.	IId	9	2.00
LIMESTONE	NO	YES	40.0	30.	IVa	7	2.00
LIMESTONE	NO	YES	25.0	30.	IVb	19	2.00
CYCLIC	NO	NO	15.0	20.	V	14	2.00
LIMESTONE	NO	YES	25.0	20.	VIa	6	2.00
LIMESTONE	NO	YES	20.0	30.	VIb	19	2.00
ALLUVIAL	***YES***	***NO***	***10.0***	***2.***	***Ia***	***2***	***1.00***
ALLUVIAL	NO	NO	10.0	20.	Ib	3	1.00
ALLUVIAL	NO	NO	27.0	9.	Ic	5	1.00
LIMESTONE	NO	YES	15.0	25.	IIa	6	1.00

WEIGHT KEY
2.00 More Confident
1.00 Less Confident

FIG. 3. An example of the parameters and resultant classes used as input for the 1st Class/Fusion expert system shell.

of studying the maps and contacting their creators in an effort to establish the thought processes used in their development. The result was a list of each output class and each of the component parameters. Also included in the list were any parameters which were implied but not explicitly stated. 1st Class/Fusion expert system shell, developed by Programs-in-Motion, was used to process the list to determine the basic Earth resource information that geological engineers use when creating a GEM (Programs-In-Motion, 1987).

Because 1st Class/Fusion is an example-based system, it permits the creation of a knowledge base in terms of various parameters, a result, and a confidence factor (Figure 3). The confidence factor reflects how much weight should be given to the specified data based on the reliability of the information supporting the class definition. To illustrate the importance of the confidence factor, Figure 3 highlights two entries with nearly identical parameters but with different weights. The weight difference relates to the methods used to create the maps. The first entry was taken from a map constructed from detailed field-tested data, while the second entry approximated large areas from sparse data. From the knowledge-base information, 1st Class/Fusion induces or creates a rule that is then displayed in a decision tree format (Figure 4a). An advantage of the 1st Class/Fusion decision tree evaluation method is its ability to optimize the rule and incorporate only those parameters that are necessary for the creation of unique classes (Usery *et al.*, 1988b). Parameters that do not affect the result are eliminated and the remaining parameters are then arranged in a sequence that will ask the fewest number of questions. For example, Figure 4a shows that only the flooding and plasticity index parameters are necessary to define Class 2.

Although the 1st Class/Fusion system determines the fewest number of questions to ask, it does not minimize the number of parameters used. Therefore, the knowledge base was manually optimized by removing a single parameter from consideration, then reconstructing the rule. Once each individual parameter had been deactivated, the next step involved deactivating two parameters from consideration, which continued until all combinations were examined. The result of this optimization effort was the identification of percent slope, presence of karst, plasticty index, and flooding potential as the basic Earth resource parameters needed for GEM production.

Once the basic Earth resource parameters were identified, the determination of an appropriate classification scheme was necessary. The classification scheme was developed using an iterative process based on the parameters determined during the deactivation analysis of the 1st Class/Fusion knowledge base, numerical values calculated by 1st Class/Fusion, and some general heuristics. An example of a heuristic used in the classification scheme is to set an upper limit of 40 for the plasticity index because soils exhibiting this characteristic behave unpredictably. The classification scheme and heuristics were validated by producing a new decision tree to see if the classes generated were still unique. If they did not prove to be unique, the process of determining the classification scheme with heuristics was repeated. The result of this procedure was a final decision tree that was then converted to if-then structured rules (Figure 4b). Once the final decision tree was developed and converted to if-then structured rules, it was possible to begin development of the GoldWorks knowledge base.

```
---- start of rule ----                                   KBGIS
FLOODING ??                                               CLASS
  FREQUENT: PI ??
          < 15.00 : ------------------------------------- 2
          > 15.00 : ------------------------------------- 1
  NO : KARST ??
     YES : PI ??
          < 22.50 : SLOPE ??
                  < 35.00 : ----------------------------- 6
                  > 35.00 : ----------------------------- 8
          > 22.50 : SLOPE ??
                  < 14.50 : PI ??
                          < 44.00 : --------------------- 7
                          > 44.00 : --------------------- 16
                  > 14.50 : PI ??
                          < 28.50 : --------------------- 17
                          > 28.50 : --------------------- 9
     NO : PI ??
          < 22.50 : SLOPE ??
                  < 30.00 : PI ??
                          < 12.50 : --------------------- 3
                          > 12.50 : --------------------- 13
                  > 30.00 : ----------------------------- 12
          > 22.50: SLOPE ??
                  < 7.00 : ------------------------------ 15
                  > 7.00 : PI ??
                          < 29.50 : PI ??
                                    < 26.00 : --- 11
                                    > 26.00 : --- 5
                          > 29.50 : --------------------- 10
OCCASIONAL : PI ??
     < 15.00 : ------------------------------------------ 27
     > 15.00 : ------------------------------------------ 4
---- end of rule ----
```

```
IF Flooding is Frequent
  and
    Plasticity is > 15.00
THEN Class is 1

IF Flooding is No
  and
    Karst is Yes
  and
    Plasticity is > 22.50 and < 44.00
  and
    Slope is < 14.50
Then Class is 7

IF Flooding is No
  and
    Karst is No
  and
    Plasticity is > 22.50 and < 26.00
  and
    Slope is > 7.00
Then Class is 11

•

•
```

a) Decision Tree **b) If-Then Rules**

FIG. 4. Translation of a 1st Class/Fusion decision tree into if-then structured rules.

Implementation of the KBGIS required selection of an appropriate GIS and expert system shell. The Earth Resources Data Analysis System (ERDAS) software package, a raster-based system, was chosen as the GIS (ERDAS, 1987), and the GoldWorks expert system shell package from Gold Hill was used to develop the expert system (Gold Hill, 1987). GoldWorks provides "a knowledge-based expert system development environment integrated with Gold Hill's Golden Common LISP (GCLISP) Developer software" (Gold Hill, 1987). It supports two different interfaces which, in turn, allow both the programmer and non-programmer to easily develop expert systems. Non-programmers can use the Menu interface, with its completely menu-driven environment, to quickly and easily develop and modify knowledge bases. This ability is of particular importance to the development and maintenance of the KBGIS. GoldWorks also supports the needs of more experienced programmers by offering the opportunity to use the Developers interface. This interface uses the GMACS editor and GCLISP which permit the creation of more powerful and specific expert systems.

GoldWorks provides easily understood and used structures for both facts and rules. GoldWorks uses a frame-based structure to represent the facts of the knowledge base. A frame, representing the structure of a class of objects, consists of a set of slots and each slot represents a potential attribute for that frame (Figures 5a and 5b). An actual occurrence of a specific object is represented by an "instance" of the frame (Figures 5c and 5d).

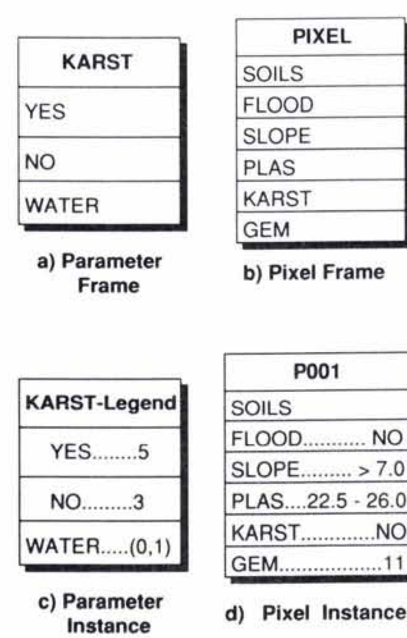


FIG. 5. Two of the frames and associated instances used in the GoldWorks KBGIS.

In GoldWorks, rules are structured in a simple if-then format and can, therefore, be easily input and modified in the Menu interface. After the rules are input into the system, they are then used to deduce new facts from the existing fact base. This process is controlled by the inference engine. The inference engine within GoldWorks supports three inference techniques; forward-chaining, backward-chaining, and goal-directed forward-chaining. The rule base can use any one, or all, of these techniques in order to deduce new facts. In the KBGIS rule base, the forward-chaining technique was used.

Only two GIS processing functions were needed to produce a GEM: recode and matrix. The recode operation allows classes to be regrouped into new categories. In other words, it allows multiple class values to be mapped into a single value. The matrix operation combines the classes of two files to produce new classes. The GIS input and output files are in a numeric pixel format but the KBGIS uses conceptual values or strings for its processing. These KBGIS conceptual values represent the symbolic pixel values of the various overlays. For example, when a karst input pixel is being processed by the expert system, it will have one of the following conceptual values: "no," "yes," or "water." These conceptual values are then used by the rule base to determine a conceptual GEM classification value, like Class II.

The KBGIS can currently produce a GEM given one of two different types of input. The first type of input requires that the basic parameters be input as separate GIS files: flooding, slope, plasticity, and karst. In addition to supplying these input files, the user must also supply the information necessary to map the input numeric pixel value (e.g., 5) to a conceptual value (e.g., "Yes") for each of the parameter files. This information is used by the system to create a legend instance for each of the parameters (Figure 5c). This instance is then used by the system to perform the input numerical to internal conceptual conversion. After all of the parameter legend instances have been created, processing of the GIS files begins. Processing consists of creating a pixel instance and then filling in the appropriate conceptual parameter values (Figure 5d). The rule base is then used to perform the matrix operation and fill the GEM slot of the pixel instance with a conceptual value. This conceptual value is then recoded into a numeric value and output to the GIS-formatted GEM file. Once all of the pixels have been processed, an ERDAS GIS trailer file is created which contains a legend to aid in interpreting the GEM.

The second type of input that the KBGIS can accept is that of a single agricultural soils file, compiled by the Soil Conservation Service (SCS). Along with the information on the general soil type found in each class, basic engineering properties are presented in a tabular format; both are used to produce a GEM file. In order to use a single GIS soils file, however, the user must first describe each possible soil type in terms of the flooding, slope, plasticity, and karst characteristics. In other words, a specific pixel instance is created for each of the possible soil types, and the appropriate conceptual attribute values are placed in the slots. The system then performs a matrix operation by utilizing the rule base, and fills the GEM slot for each of these instances. To produce a GEM file now simple requires a KBGIS recode of the soils file using the pixel instances, and the creation of a trailer file.

Testing of the KBGIS was done by selecting an area with a preexisting manually produced GEM and generating a KBGIS GEM for it. These two maps were compared and the results of the comparison analyzed. It was decided that a single agricultural soils file would be used for the GEM generation. Therefore, the published map of the agricultural soils information for the Creve Coeur quadrangle (Anon, 1982) was digitized using ERDAS. It was then input into the KBGIS and a GEM was produced. This GEM was displayed side by side on a graphics system with the digitized manually produced GEM. It was concluded that a "good" correlation existed between the two maps, taking into account land-use changes that may have occurred in the 10 years since the manual GEM was produced (Usery *et al.*, 1988b). The major discrepancies were caused by differences in the classification methods, rather than problems with the system (Usery *et al.*, 1988b). In fact, the KBGIS produced a more detailed GEM than the one produced manually. It was therefore concluded that a KBGIS could be used to produce an accurate GEM.

Unfortunately, it was also concluded that the execution time required by the KBGIS severely hampered its practicality as a production system. For example, production of the Creve Coeur GEM took in excess of 76 hours. This meant that, although the KBGIS automated the GEM production process and cured the weaknesses of the manual production system, it failed as a time efficient production system (Cress, 1989).

DEVELOPMENT OF THE CONVERSION EXPERT SYSTEM

Testing of the final GoldWorks KBGIS supported the theory that an automated approach could be used to produce GEM maps. In fact, automation strengthened the production process by standardizing the classification definitions, reducing the time commitment required of an engineer, and removing the need to start from scratch each time a specific site selection GEM was required. Unfortunately, the GoldWorks implementation still required a significant amount of production time. In response to this problem, an independent research project was conducted to determine if a more efficient method of implementation could be found. This project consisted of, first, identifying all the characteristics of the KBGIS as well as the primary weaknesses of the current implementation method, and second, developing a more efficient KBGIS production system.

The main component of the KBGIS is its application knowledge base, which consists of the specific geological engineering rules needed to generate the GEM from the basic Earth resource parameters. As mentioned earlier, this knowledge base consists of two parts: data structures for the basic Earth resource parameters and the rules which define the GEM classifications. The data structure used for each of the parameters is a frame consisting of a set of slots which represent the valid conceptual values (Figure 5a). The rules follow a fairly standard if-then format: the "if" or antecedent consists of a combination of parameters and their associated conceptual values; the "then" or consequence sets the classification value. An important characteristic of these rules is that they are, and will always be, mutually exclusive because classification schemes demand that unique parameter combinations be established for each classification definition.

Even though the general characteristics of the application rule base are static, the specific rules themselves are dynamic. This stems from the regional approach used for the development of the KBGIS rule base. It was determined early in the development process that it is impractical to develop a single set of basic Earth resource parameters and associated rules for use with all GEM production systems. As a result, any KBGIS development system must permit the easy development and modification of the rule base. Also, because a geological engineer would be the primary KBGIS rule base developer, a user-friendly environment is important. Because the GoldWorks Menu interface provides such an environment, a final characteristic of the KGBIS is that its rule base be developed in GoldWorks.

Although the GoldWorks environment is obviously vital to the development of the application knowledge base, it is inefficient when used to implement the resultant knowledge base in a production environment. A primary weakness of the KBGIS GoldWorks implementation is its inability to efficiently process large data files. Although excessive execution time could be avoided during the development process by using limited test data sets for preliminary prototyping and testing, production systems would be unable to avoid this problem. This is especially true because a typical GIS layer could easily contain 250,000 or more bytes of data. Another weakness of the KBGIS as a production system is its reliance on access to a GoldWorks package for implementation. Even after the final KBGIS rule base has been developed, any use of the production system would have to occur in the GoldWorks environment. This could be an inhibiting factor for a smaller GEM production environment, such as a field office, because licensing multiple GoldWorks packages for use as production systems would require a substantial financial investment. Unfortunately, this could eventually result in only larger GEM production facilities adopting the KBGIS production system, which would defeat the goal of standardizing GEM classifications.

On the other hand, because of the exclusive nature of the rules, implementation in an expert system shell is not a requirement. Once the rule base is established for a given region, it can be converted into a more traditional programming language for implementation. This not only relieves the requirement of access to GoldWorks for each of the production systems, but it also permits implementation in a language better equipped to handle large data files. Unfortunately, conversion from the GoldWorks implementation to another language would require access to a skilled computer programmer. Because of the dynamic nature of the rule base, this conversion process would have to be performed multiple times and possibly over a prolonged time period. These problems appeared to make a conversion approach impractical, until it was realized that an expert system could be developed to perform the conversion. Development of such an expert system would remove the need for a programmer because it would replace the programmer in the conversion process. In other words, a Conversion Expert System (CES) could accept as input any GoldWorks KBGIS and convert it into another computer language. This new program, after compilation, could then be used as the GEM production system.

The first step in the development of a CES was to choose an appropriate language into which the GoldWorks knowledge base would be converted. The next step was to confirm that this conversion would result in significant time savings for the production system. FORTRAN was chosen as the conversion language because of its inherently fast execution time for data intense operations. Another factor in favor of FORTRAN was its availability. Because the majority of the computer systems already support FORTRAN, additional software purchases would generally not be needed in order to utilize the KBGIS FORTRAN production system. Once FORTRAN was chosen, the existing GoldWorks rule base for the Midwestern region was manually converted into equivalent FORTRAN code to establish that such a conversion would result in a more efficient implementation. A sample ERDAS data set was used to develop a GEM, using both the FORTRAN program and the KBGIS expert system. Comparisons of the two GEMs confirmed that the GEM produced by the FORTRAN system was identical to the GoldWorks GEM. A comparison was also made of the execution times of the two GEM production systems. Execution was performed on the same hardware in order to eliminate any hardware-related execution factors. The FORTRAN system produced a GEM in under six minutes, whereas the GoldWorks system required over three hours. It is apparent that the FORTRAN production system significantly improved the efficiency of the KBGIS rule base.

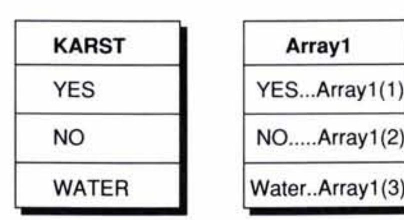
Once the FORTRAN KBGIS approach had been validated, an automated system to convert from GoldWorks to FORTRAN was needed. The GoldWorks expert system shell was again chosen as the development tool for the CES, this time because it allowed the creation and use of powerful LISP functions in its Developers interface.

The first task of this expert system was to convert the GoldWorks frame-based data structures into appropriate FORTRAN data structures. FORTRAN arrays were used because they allow a one-to-one mapping of KBGIS frames and slots to FORTRAN arrays and elements. Conversion of a frame into an equivalent FORTRAN array was done by creating an instance of the specific frame, and then giving the name of the equivalent FORTRAN array element as a value for each of the slots (Figure 6a and 6b). The name of the array is used as the instance name, and a key list containing the frame name and associated array name was constructed (Figure 6c). An analogous FORTRAN array data structure was created for each of the frames which represent one of the basic parameters. A search of the rule base determined which frames to convert; a list was generated of the parameters found in the antecedent of the rules which set the GEM classification. This parameter list was then used as a key for indicating those frames needing conversion. The rules were converted after all of the appropriate FORTRAN data structures were created.

Only the GoldWorks rules used to deduce a GEM classification needed to be converted into FORTRAN. Identification of these rules was done by retrieving the consequence of each rule and searching it for the name used to represent the class slot of the pixel frame. "GEM" is the slot name in the existing KBGIS. If a match was found, then the specific classification value set by the rule was checked for validity. Confirmation of validity was done by making sure that the classification value specified in the rule matched one of the slots for the "GEM" frame in the KBGIS. Once it was established that a rule needed to be converted, a list of the rule's antecedent was retrieved. This list can consist of several layers, depending on the complexity of the antecedent. Each layer of the list was searched until a minimal sublist was found. A minimal sublist is a sublist which contains the pattern used to search the rule base, i.e., it contains a parameter name and associated conceptual value. The minimal sublist was then converted into FORTRAN code. If an antecedent list contained multiple minimal sublists, then the appropriate logical connective ("or," "and") was also converted into FORTRAN. This process was recursively repeated until all of the minimal sublists in the rule's antecedent were converted. Once the rule's antecedent had been processed, the rule's consequence was retrieved and converted. This conversion was much simpler because the list of the consequence contained only the GEM slot name and an associated conceptual GEM classification.

The only remaining part of the KBGIS system to convert is the user interface, which was relatively simple. The only anticipated changes are the names and number of the basic Earth resource parameters based on the region covered by the knowledge base. The formats of both the input and output files remain constant because they are based on ERDAS file format.

To use the CES, the KBGIS rule base must first be loaded into



FRAME	ARRAY
KARST	ARRAY1
SLOPE	ARRAY2
FLOODING	ARRAY3
PLASTICITY	ARRAY4

c) Key List

FIG. 6. Conversion of a GoldWorks frame structure into a FORTRAN array.

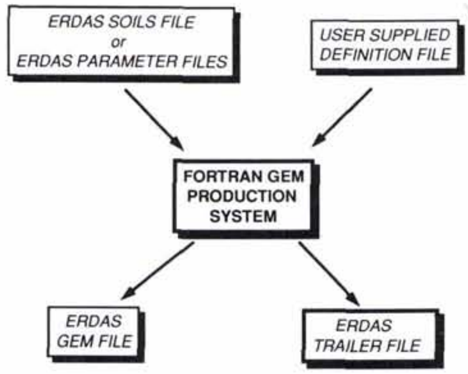


FIG. 7. FORTRAN-based KBGIS GEM production system.

the GoldWorks environment. Once this is done, the CES can be executed. The CES is extremely easy to use and requires a minimum amount of input from the user. Furthermore, input from the user is not only user correctable but also checked for validity by the system. If either the user or the system decides that an incorrect entry has been made, the user is given the opportunity to enter the input again.

The opening screen of the CES gives the user the option of exiting the system or continuing. After entering the system from the opening screen, the system requests the directory path name for the output FORTRAN code. This consists of six FORTRAN files, called ftr0.dat to ftr5.dat, that must be concatenated together in ascending numerical order to produce the complete FORTRAN source code. This can be done using the DOS copy command. The user is then asked for input and output variable names, which are arbitrary, as well as a default conceptual class value, which must be a valid GEM classification. The variable names are used in the FORTRAN code to reference the input and output pixel values. And the default class value is used as the output class value when none of the other GEM rules apply. Finally, the user is asked whether the specified parameter, which the system has retrieved from the KBGIS rule base, can be derived from a single agricultural soils file or is present as an independent overlay. The system will allow both formats to be used in a single application.

Once the generated FORTRAN code has been compiled and linked into an executable program, it can be used as a GEM production system (Figure 7). This system is equivalent to the KBGIS system in terms of input and output, i.e., it requires one of two types of ERDAS input and produces two output files (the GEM and a trailer) in ERDAS format. The only difference between the systems is their methods of processing; the FORTRAN program performs numeric to numeric recodes instead of numeric to conceptual to numeric recodes like the KBGIS. In addition, the FORTRAN code is limited to mapping single values to single values, unlike the KBGIS which supports multiple values to a single value mapping. This is, however, the only additional constraint added by the FORTRAN production system. If a multiple to single mapping is required of the system, it can be performed in ERDAS as a pre-processing step.

CONCLUSION

The CES was tested by converting the Midwestern KBGIS rule base into FORTRAN. After the FORTRAN program was compiled and linked, it was then used to produce a GEM. This product was compared to a GEM produced by the KBGIS from the same ERDAS input. Test results showed no differences between the two GEMs. Comparisons also were made of the execution times of the two production systems. The FORTRAN system processed 179 bytes per second (0.0056 sec/byte) as compared to the KBGIS which processed only 0.77 bytes per second (1.30 sec/byte). This comparison was based on a small data set of 62,500 bytes. Furthermore, the time required to process a byte grows in the KBGIS GoldWorks system proportional to file size; the FORTRAN code is independent of file size, so increasing the data set to production sizes would result in additional time savings. It is apparent that the FORTRAN production system performs in a significantly more efficient manner.

It can be seen that the KBGIS approach to development of a GEM production system has definite benefits, but fails as a production system due to the time elements involved. The introduction of an intermediate development step, however, was able to bridge the gap between ease of development and efficiency of use. This bridge is the Conversion Expert System. This system significantly improves the practicality of the automated production of GEMs by drastically reducing the production time. The FORTRAN approach was also able to solve the possibly inhibiting effect of the cost of the system. Although expert system development tools are still a requirement to develop a KBGIS, they are not necessary to use the production system at each field site. Once the KBGIS has been developed and converted into FORTRAN, any system that can execute FORTRAN can be used as a production system. This means that a hub approach could be taken by a production company with potentially significant financial savings. A hub approach means that only the central regional office would need to purchase the expert system shell package and associated hardware. The KBGIS could then be developed and converted at the central office, and only the resulting executable program would need to be sent to the field offices.

In conclusion, the development of a KBGIS for the production of GEMs resulted in major improvements over existing production methods. The development of a Conversion Expert System, used in association with the KBGIS, made the resulting production process both efficient and affordable.

REFERENCES

Anon, A., 1982. *Soil Survey of St. Louis County and St. Louis City, Missouri*, United States Department of Agriculture, Soil Conservation Service, 137 p.

Cress, J. J., 1989. *Development of an Expert System to Convert Knowledge-Based Geological Engineering Systems into Fortran*, University of Missouri-Rolla, Unpublished Master of Science Thesis, 28 p.

ERDAS, Inc., 1987. *ERDAS-PC and PC Kit: Image Processing System User's Guide*, ERDAS, Inc., 2801 Buford Highway, Suite 300, Atlanta, Georgia.

Gold Hill, 1987. *GoldWorks Expert System User's Guide*, Gold Hill Computers, Inc., 26 Landsdowne Street, Cambridge, Massachusetts.

Lutzen, E. E., and J. D. Rockaway, Jr., 1970. *Engineering Geology of the Creve Coeur Quandrangle, St. Louis County, Missouri*, Engineering Geology Series No. 2, Missouri Geological Survey and Water Resources, Rolla, Missouri, 19 p.

———, 1971. *Engineering Geology of St. Louis County, Missouri*, Engineering Geology Series No. 4, Missouri Geological Survey and Water Resources, Rolla, Missouri, 23 p.

Programs-In-Motion, 1987. *1st Class/Fusion User's Guide*, Programs-In-Motion, Inc., 10 Sycamore Road, Wayland, Massachusetts.

Smith, T. R., and M. Pazner, 1984. Knowledge-Based Control of Search and Learning in a Large-Scale GIS, *Proceedings, International Symposium on Spatial Data Handling*, Zurich, Switzerland, pp. 498–519.

Usery, E. L., R. R. P. Deister, and D. J. Barr, 1988a. A Geological Engineering Application of a Knowledge-Based Geographic Information System, *Proceedings, American Congress on Surveying and Mapping/American Society for Photogrammetry and Remote Sensing*, Vol. 2, Cartography, St. Louis, Missouri, pp. 176–185.

Usery, E. L., P. S. Altheide, R. R. P. Deister, and D.J. Barr, 1988b. Knowledge-Based GIS Techniques Applied to Geological Engineering, *Photogrammetric Engineering & Remote Sensing*, Vol. 54, No. 11, pp. 1623–1628.