Incorporating A New Computational Reasoning Approach to Spatial Modelling

Alec Holt

Spatial Information Research Centre, Department of Information Science,
University of Otago, P.O. Box 56, Dunedin, New Zealand.
Aholt@commerce.otago.ac.nz, Phone : +064 03 479 8301. Fax +064 03 479 8311

Decision support systems, statistics and expert systems were some of the mainstay techniques used for modelling environmental phenomena. Now modelling systems utilise artificial intelligence (AI) techniques for the extra computational analysis they provide. Whilst operating in a toolbox environment and by adopting AI techniques, the geographic information system (GIS) modellers have greater options available for solving problems. This paper outlines a new approach in applying artificial intelligence techniques to solve spatial problems. The approach combines case-based reasoning (CBR) with geographic information systems and allows both techniques to be applied to solve spatial problems. More specifically this paper examines techniques applied to the problem of soil classification. Spatial cases are defined and analysed using the case-based reasoning techniques of retrieve, reuse, revise and retain. Once the structure of cases are defined a case base is compiled. When the case base is of sufficient size, the problem of soil classification is tested using this new approach. The problem is solved by searching the case base for another spatial phenomena similar to *that which exists. Then the knowledge from that searched case is used to formulate an answer to the problem. A comparison of the results obtained by this approach and a traditional method of soil classification is then undertaken. This paper also documents the saving data concept in translating from decision trees to CBR. The logistics of the problems that are characteristic of case-based reasoning systems are discussed, for example, how should the spatial domain of an environmental phenomena be best represented in a case base? What are the constraints of CBR, what data are lost, and what functions are gained? Finally, the following question is posed: "to what real world level can the environment be modelled using GIS and case-based reasoning techniques"?*

Acknowledgements:

- Chris Price and the University of Wales for providing the CBR software called CASPIAN.
- Alan Hewitt and Manaaki Whenua -Landcare Research (Dunedin, New Zealand) for providing expert knowledge and data.

INTRODUCTION

Geographic information systems (GIS) are progressing towards systems which incorporate greater geocomputational functions. Three factors provide impetus behind this progression. First, spatial problems are inherently difficult to solve and the spatial information and modelling communities recognise that the lack of analytical and modelling functionality is a major deficiency of current GIS (Fischer & Nijkamp 1993). Second, as GIS databases mature users seek techniques which allow for further analysis (Burrough & Frank 1995). The third, which is a new concept, suggests that with the evolution of mature databases the vendors or data owners will move to develop applications and tools for their clients (Benwell 1996). These factors impel GIS progression towards a toolbox environment.

In addition GIS's progression towards a toolbox environment can be explained by its position as a platform for integrating various databases and systems. Decision support systems, expert systems, neural networks, fuzzy logic and connectionist systems, for example, are some of the many databases and systems which have been successfully coupled with a GIS. AI is also a good integrator and offers another avenue for providing geocomputation features. The various hybrids currently being implemented provide evidence of this. In a GIS-AI hybrid GIS could be used to represent and display the problem and solution while the AI techniques could be used to process the bulk of the problem solving. Some modelling systems use GIS or graphical display embedded in other platforms, including delphi and visual basic applications. An AI hybrid in comparison is also an excellent platform as some have graphical display techniques built into their systems. This is particularly applicable to medical applications.

As a result of these three factors a number of research paths can be taken to provide further geocomputational functions. In observing the diverse subjects of recent GIS conferences, analytical and geocomputation techniques are prolific. This paper identifies overlapping characteristics and diverse functions available in the following list of disciplines; statistics, cognitive science, knowledge acquisition, databases, case-based reasoning (CBR), inductive learning and knowledge discovery. In highlighting these diverse disciplines it will be seen that GIS modellers could progress their GIS systems to greater geocomputation levels by adopting some of the above techniques.

It is suggested that if CBR is incorporated with GIS it will further the geocomputational level of the GIS. CBR offers this GIS-AI Hybrid software reasoning from data, explanation, adaptation, extended generalisation techniques, inference making abilities, constraining a search to the solution template, generation, refinement, validation and maintenance of knowledge bases. These features help in planning, forecasting, diagnosis, design, decision making, problem solving and interpretation.

This paper will focus on these added features CBR offers the spatial reasoning system (SRS) (Holt & Benwell 1995a). With the potential of these added features, CBR aids a GIS and also suggests a new method for modelling spatial data.

Other research that has furthered the GIS progression towards a more geocomputational system include rule and knowledge-based approaches (Webster 1990; Smith & Yiang 1991; Skidmore *et al.* 1991), hybrid connection systems (Kasabov and Trifonov 1993), multiple criteria decision-making methods (Jankowski 1995) and a more innovative research approach where spatial reasoning is used to identify a given situation with other known typical scenarios (Williams 1995). These different analytical approaches are being coupled to form soft computing, for example neural networks with expert systems (Skidmore, *et al.* 1991). Case-based reasoning has been coupled with decision support systems (Burstein and Smith 1994). An interesting and important connection has seen the integration of case-based reasoning and neural networks. This method proposes a co-processing hybrid model for classification by coupling case-based reasoning and neural networks (Malek & Labbi 1995). Interest in hybrid systems is beneficial as effectual new systems for example, combinations such as neuro-fuzzy systems use the strengths of both neural networks and fuzzy systems to provide a more intelligent system. In GIS and modelling communities advances in AI systems through various combinations, bode well for strengthening the analytical capability of a GIS.

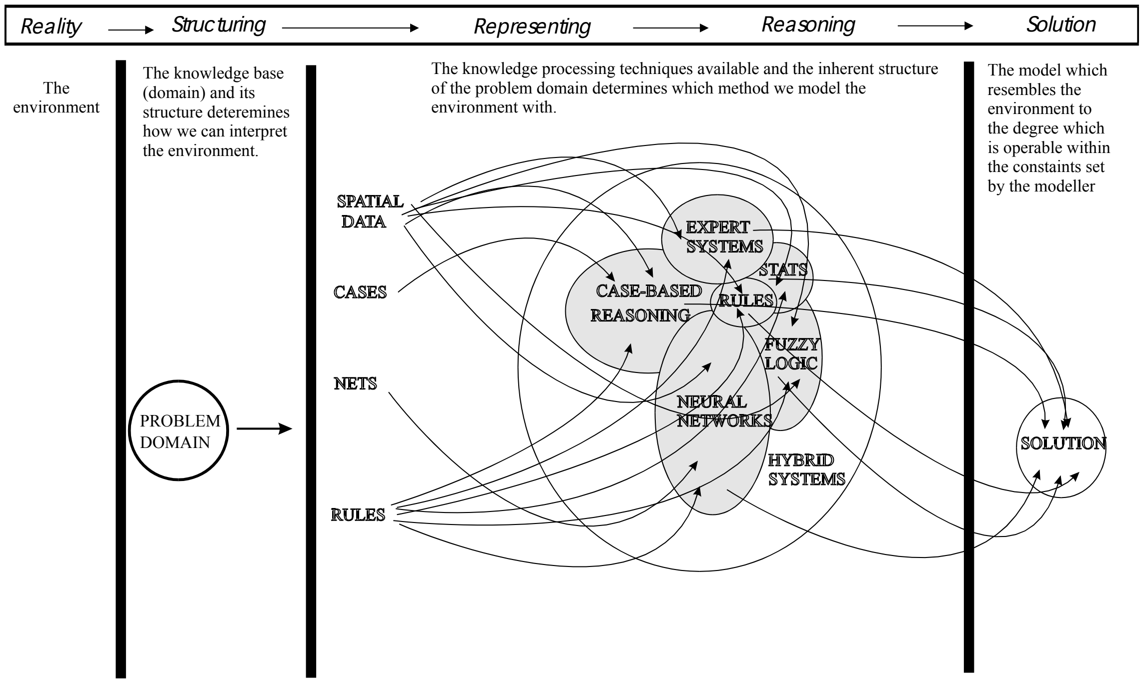
New techniques are illustrated and discussed in this paper. In order to determine “which techniques best suit which applications” an understanding of how the problem data can be best represented, and which of these working units the problem domain allows is necessary. If many working units are allowed the question of which technique should be adopted remains. A difference exists between the tasks AI techniques perform and those which can be best applied to spatial phenomena. Some of the techniques offered by AI techniques may not be required by the spatial modelling community. This paper focuses on CBR as one of these new tasks and documents the saving data concept in moving from decision trees to CBR. Which techniques best suit which problems? The following are limiting factors as specific data can be represented in certain ways and some techniques, such as GIS, in effect force data to be stored in a particular manner. The manner in which data is represented in effect dictates which analysis or reasoning techniques could be used. To reduce the impact of these limiting factors one could employ a large platform base. Having an AI-Hybrid, for example, would allow a number of analysis functions to be applied through one representation technique. A summary table of available techniques, their use, and for what type of spatial phenomena are they particularly suited (some techniques may be available but they are not chosen to be used) is currently under development (Holt in press). This has given rise to the notion of intelligent spatial information systems (Leung 1993; Laurini & Thompson 1992), which have adopted some, if not all, of the following statistical and analytical approaches to be discussed.

AI PROBLEM SOLVING PATHWAYS

Not all AI techniques are being fully utilised in the spatial information systems realm. AI usage for spatial problem solving has tended to be *low level processing, f*or example, in the classification of image patterns, primarily to complete images and to clean noisy data (Openshaw 1993; Kasabov & Trifonov 1993). AI techniques in comparison should be used to provide better decision support and more intelligent modelling systems. These systems could be used to solve spatial dilemmas which current GIS's fail to do. Problems requiring further analytical processing could be specifically targeted by these systems. The GIS-AI hybrid provides *high level processing* and, therefore, increases the analytical processing ability of a GIS. The ability to reason may produce this higher level of processing. Therefore, it is proposed that a GIS-AI hybrid with the ability to reason should be developed (Holt and Benwell 1995a).

AI techniques available include, fuzzy logic (FL), neural networks (NN), case-based reasoning (CBR), genetic algorithms (GA), statistical models, knowledge and rule-based, fragmentation indexes and hybrid connection systems. A variety of pathways are therefore available to solve complex problems. An approximate map of the pathways of computational methods available for analysing spatial data has been drawn (figure 1). More than one AI pathway (some have overlapping functionality) could be used to solve a problem as each solves the problem differently according to their encapsulated functionality. One such pathway, which indicates how to solve complex spatial problems using a proposed GIS-CBR hybrid, is indicated in figure 1. In the course of the evolution of the spatial analytical toolbox it may be possible that a different GIS-hybrid is formed. Some AI techniques can increase their level of their performance if they are combined with other AI-techniques to provide hybrid systems. CBR-NN, FL-NN, for example, become more effective when combined.

Figure 1. Problem solving pathways



This may provide a comprehensive GIS where connection hybrids may become integrated with GIS. The GIS-CBR hybrid, which will be used to further the comprehensive nature of GIS, will be discussed. Two examples using the reasoning of a CBR to manipulate spatial data will be illustrated. The examples use CBR to evaluate test sites with previous spatial sites and amalgamate the spatial similarities of the test and previous sites to provide decision support to solve the spatial dilemma.

CASE-BASED REASONING

CBR is a general paradigm for reasoning from experience. It assumes a memory model for representing, indexing and organising past cases and a process model for retrieving and modifying old cases and assimilating new ones (Kolodner 1993).

It is important to define a case as they form the basic elements of CBR systems.

> *A case is a contextualised piece of knowledge representing an experience that teaches a lesson fundamental to achieving the goals of a reasoner (*Kolodner 1993:13).

A case is a problem-solution pair. This emphasises the problem solving mechanism of CBR using the problem-solution pair to solve a similar problem. The two components of the pair are input and stored cases. *Input cases* are descriptions of specific problem situations. *Stored cases* encapsulate previous specific problem situations with solutions and outcomes. *Stored cases* contain a lesson and a specific context in which the lesson is applied. The context is used to determine when the lesson may apply again. These input cases are used to find other similar situations in a spatial manner (Leake 1995; Aha 1994).

Cases are examples which have occurred in reality or problems that have occurred and been solved (success and failures) by a problem solving mechanism. (Althoff *et al.* 1994)

The major components of a case include;

1. **Problem/situation description**: the state of the real world at the time the case was happening and, if appropriate, what problem needed to be solved at that time.
2. **Solution**: the stated or derived solution to the problem specified in the description or the reaction to its situation.
3. **Outcome**: the resulting state of the world when the solution was carried out.
4. **Extensions**: the context (justification) which links to other cases and the failures encountered.

(Kolodner 1993; Althoff *et al.* 1994)

These components of a case are the cogs of the case-based reasoning-cycle, or the solution and outcome components which make it possible to reuse, revise and retain cases.

More specifically, case-based reasoning is defined as;

> *a cyclical artificial intelligence problem solving paradigm that stresses reuse of solutions to similar problems, where solutions are maintained in a carefully indexed memory* (Aha 1994:3).

The above definition suggests that the components of CBR are representation, indexing and storing of cases for problem solving by retrieving, adapting, explaining, critiquing and the interpreting of previous situations. This process is used to create a solution to a problem using previous information. It is suggested that these components be added to GIS to complement its analytical functionality to build a *spatial reasoning system*. The proposed *spatial reasoning system* is designed to test the hypothesis that case-based *reasoning can be used to complement spatial analysis in GIS.*

Tasks that all CBR methods undertake include identifying the current problem situation, finding a past case similar to the new one, using that case to suggest a solution to the current problem, evaluating the proposed solution and updating the system by learning from experience.

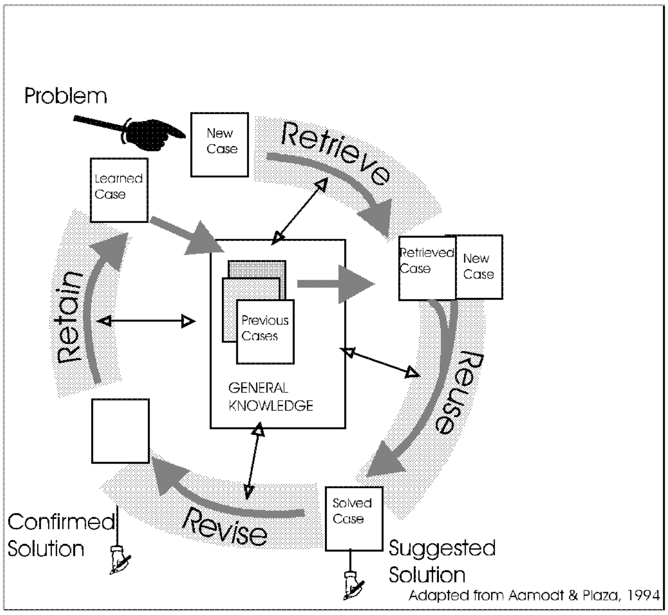
THE CASE-BASED REASONING-CYCLE

A CBR cycle may be described by the following four processes see figure 2 (Aamodt & Plaza, 1994; Kolodner 1993; Aha 1994; Leake 1995);

1. **retrieve** the most similar case(s),
2. **reuse** the information and knowledge in that case to solve the problem,
3. **revise** the proposed solution,

4. **retain** the parts of this experience likely to be useful for future problem solving.

Figure 2 CBR Cycle



CBR is a relatively new tool for solving spatial problems. There have been previous applications of CBR to solve spatial phenomena and environmental problems (Branting & Hastings 1994; Berger 1994; Jones & Roydhouse 1993; Kolodner 1993; Lekkas *et al.* 1994).

- Berger (1994) uses CBR to solve a medical problem using spatial data. The application is called *ROENTGEN* and is a CBR system that aids in planning radiation therapy for new patients based on geometrically similar previous patients.
- Branting and Hastings (1994) have developed a system called *CARMA* that uses model based reasoning and CBR to combat rangeland pests.

- Jones and Roydhouse (1993) produced a system called *MetVUW Workbench* which has been used for the retrieval and the display of historical meteorological data.
- Lekkas *et al.* (1994) developed a system called *AIRQUAP* which has been used to predict air pollution levels.
- Kolodner (1993) designed systems called ARCHIE. This is a system which aids an architect to design a new building. The cases represent knowledge about previous designs of buildings with similar specifications and situations.

THE SPATIAL REASONING SYSTEM (SRS)

Some definitions for reasoning in the GIS community include spatial cognition and the representation of knowledge (Hernandez 1993; Williams 1995). Frank (1996) defined reasoning as “the conceptualization of situations as space”. For the purpose of this research r*easoning* in this paper means the ability to reason; learning; thinking and the ability to draw on conclusions from facts (Holt 1996).

In problem solving a GIS uses raw data, not processed data, as there is no cycle and no facility to retain the solution. Therefore, there is no reuse of a previous solution or the process taken to derive that solution. This paper proposes a GIS-AI Hybrid called the spatial reasoning system. CBR offers this GIS-AI hybrid software an ability to reason, explain, adapt, extended generalisation techniques, inference making abilities, constraining a search to the solution template, generate, refine, validate and maintain knowledge bases. These features help in planning forecasting, diagnosis, design, decision making, problem solving and interpretation.

The Spatial Reasoning System (SRS) will eventually be used;

1. As a problem solving tool which has the ability to reuse previous similar spatial problems and their solutions to solve a current problem (Holt & Benwell 1996).
2. *As a problem solving tool which has the ability to reuse previous similar spatial problems and their solutions to solve a current problem*, with the added function of using a graphical interface to enter criteria.
3. As an exploratory spatial data analysis technique for data mining/trawling and pattern searching/matching (Holt in press).
4. As a new method to represent and store spatial data. Storing data as spatial cases, equivalent to object oriented languages, but having the added benefit of learning features.

With respect to the SRS this paper will describe the first option above. It is suggested that initially a GIS-CBR hybrid could be used to help in spatial decision making, spatial problem solving and spatial interpretation. This would produce a more intelligent GIS system as CBR could be used in the following ways;

1. The GIS-CBR hybrid is used *to facilitate searches* and answer the following questions. *Are there any other spatial phenomenon such as this? Identify extreme*

areas, evidence of trends, patterns or other variations. Is there clustering? If so, what attributes are associated with that phenomenon? In finding a similar spatial pattern a GIS is needed to display and store the data. CBR provides the functionality to find a similar pattern and, more importantly, to analyse its properties. These properties would extend from the obvious spatial pattern to other attributes associated with that spatial pattern. This type of functionality could be used for classification or in solving more complex problems using previous experiences.

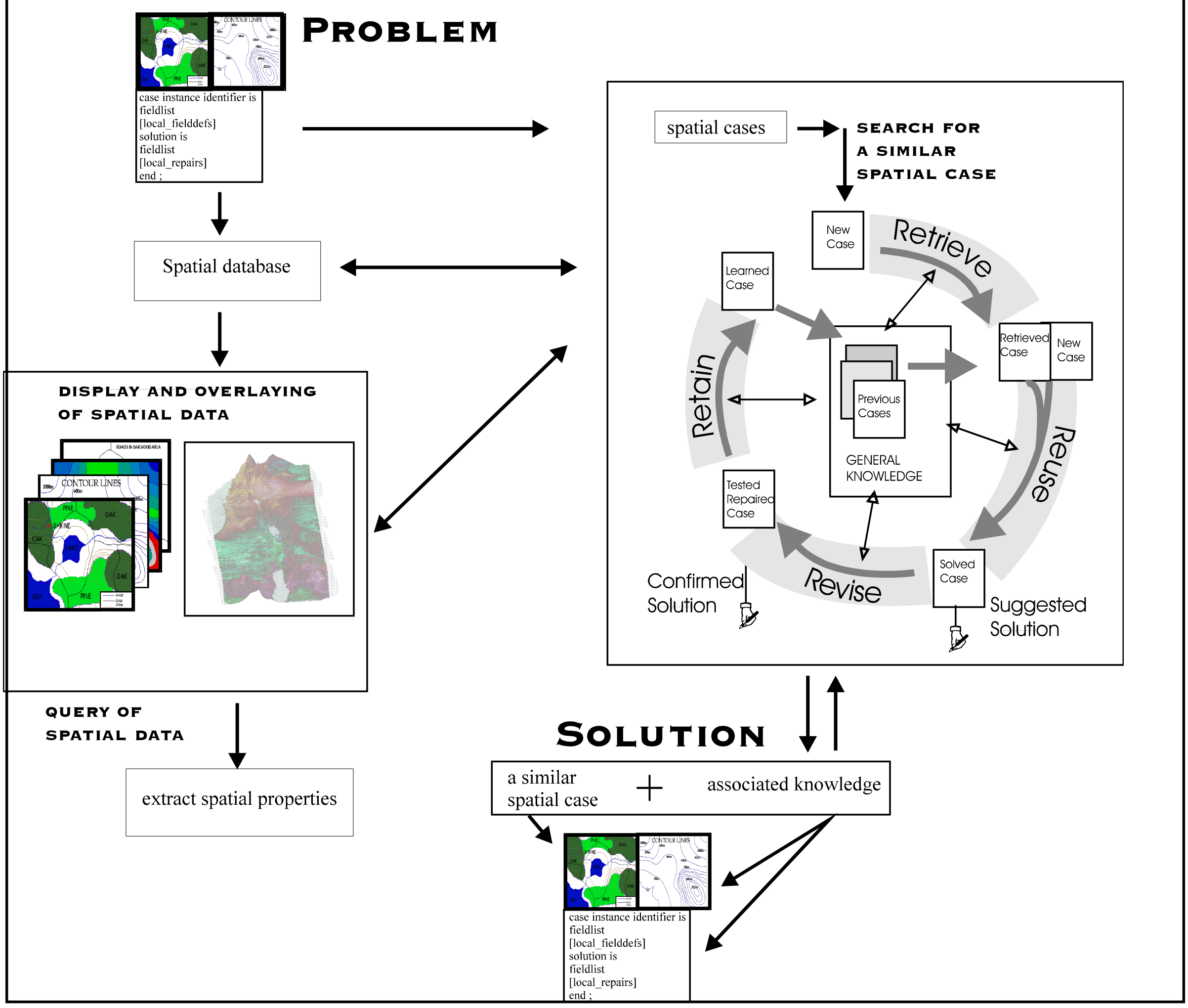
2. To make simulation possible. This is useful for the estimation and prediction of spatial phenomena including the *display of spatial-case distributional properties.*

3. Providing new opportunities in spatial analysis via information retrieval and pattern recognition. The following questions may be answered; Is there evidence of clustering in respect of specified sources or possible causes? What spatial associations exist between cases? Would a GIS-CBR model describe spatial relationships better?

4. The GIS-CBR hybrid is used *to facilitate queries* and answer the following questions; Which spatial phenomena have the following criteria? What attributes are associated with a spatial phenomena with these criteria?

The path taken to solve a problem using the GIS-CBR hybrid is shown in figure 3.

Figure 3 A Spatial Reasoning System



These criteria have spatial properties and the benefit of using a GIS for selecting slope, height and aspect include its ease of interpreting, manipulating and representing spatial data. A fully integrated GIS-CBR hybrid would have the ability to enter spatial data directly from digital maps and digital terrain models into a CBR. Once the data are entered the select action searches for a similar case, which is then displayed with any associated attributes. CBR provides the unique function of allowing further information related to the similar case to be used. These data can be saved as new cases if a decision is made based on previous cases. This function indicates CBR's learning ability. This model has been tested and has provided satisfactory results (Holt & Benwell 1995a,b). The example of ZONATION, which uses soil to portray the spatial reasoning concept will be outlined. This application provides an interesting focus as soils have implicit spatial distribution properties.

ZONATION

As well as displaying CBR techniques in a SRS, it is also suggested that knowledge can be saved in a toolbox environment, for example, in this paper the knowledge saving transition is from decision-trees to CBR.

More specifically this research applies a combination of techniques to the problem of soil classification. The logistics of the problems that are characteristic of case-based reasoning systems are discussed. *ZONATION* employs Irvin *et al.* (1995) and Hewitt's (1995) philosophy of using landforms to aid in the classification of soil series. An experiment is conducted based on data derived from Hewitt's series of tests.

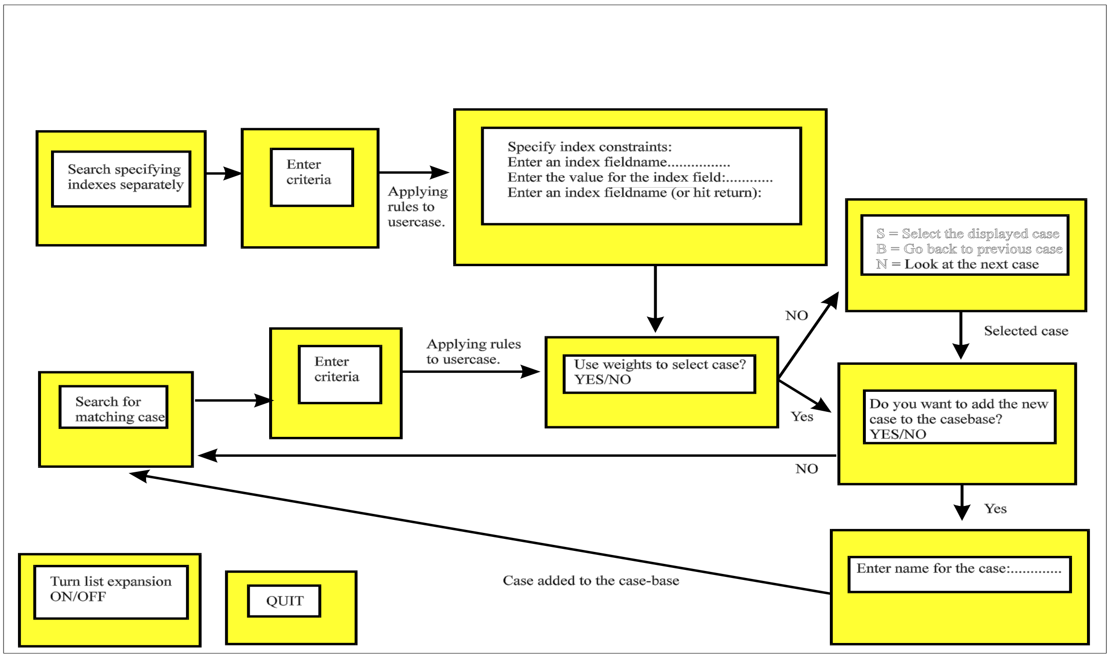
ZONATION employs a method for soil classification which utilises spatial information system techniques to classify individual pixels of a digital terrain model according to its membership in a landform class (Hewitt 1995; Irvin *et al.* 1995). These classes are determined by the natural clustering of the data in attribute space. Attributes central to this classification include, land element, slope, aspect, A_horizon texture, B_horizon texture, soil depth, A_thickness, dung, % gravel volume, % gravel weight, % carbon weight and % carbon volume (Hewitt 1995). Because these factors are also important to soil forming processes, soil classes should nest within landforms (Irvin *et al.* 1995). *ZONATION* adapts this philosophy by using the attributes and classes of Hewitt's criteria as fields and goals which are used to define case instances and, to store the attributes which are used to predict soil classification types of new zones.

The sequence for running *ZONATION* is as follows;

1. The user provides a case for comparison.
2. The program performs an index search and finds a subset of cases that match all the index constraints. The index constraints are taken from the field values provided by the user. The program searches the case base for the subset of cases that match all the index constraints exactly. ZONATION uses land_element as an index, therefore grouping all cases with the same land_element before making a selection.
3. If no cases match all the index constraints (for instances when there are only a few cases in the case base), the system prompts the user to search for different index values. If there are no cases which match all the index constraints, the user is informed and is prompted to enter new values for the index constraints. These may be made more general by specifying abstraction values or by specifying fewer constraints.

4. A case is selected from the subset. After the index search is completed the case matcher is invoked to scan the subset of cases to find the one with the highest weight value. This is selected and the repair rules are then applied.
5. Repairs are carried out on the selected case. On occasions additional information is requested after a case has been selected. Sometimes a repair rule can cause the current case to be abandoned and the selection process to begin again.

Figure 4. The Sequence For Running ZONATION



If the user is dissatisfied with the previous matching case(s) further cases may be examined. This is continued until they are satisfied with a matched case or until the user exhausts all possibilities.

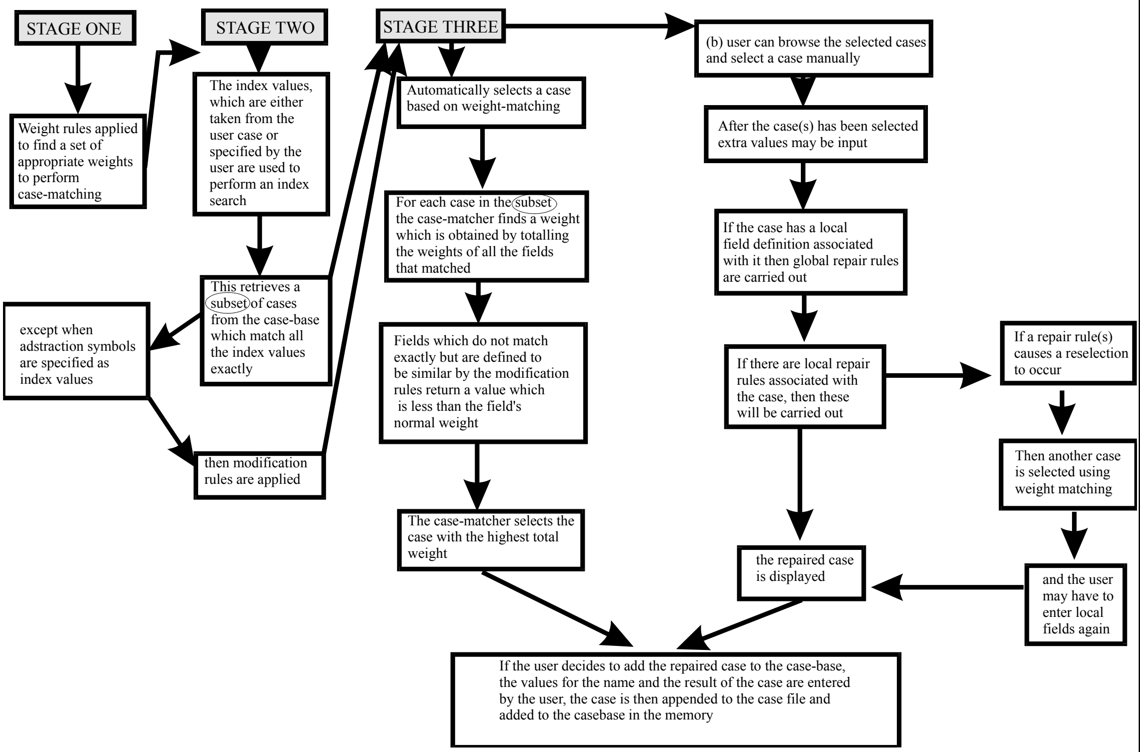
For the ZONATION case file blocks of code were written to define; *introduction, case definition, index definition, modification definition, weight rule definition, repair rule* *definition, and case instance*. The *introduction* block contains introductory text which is displayed when the program has finished checking the case file. The *case definition* sets the types and the weights of the problem fields that may appear in a case. The information in the *case definition* is used for checking input cases while the weights are used to aid the case-matching process. The *index definition* sets the fields used as indexes when searching for a matching case. A case base should have at least one field used as an index. The type of index field must be enumerated. The *weight rules definition* sets rules that may be applied to change the weights used for matching cases. The *modification definition* sets the modification rules and provides a means of specifying that certain symbols or numbers are similar. This is undertaken first for matching purposes and provides a means of specifying symbols as abstractions of others and second for making the search more general or for defining generalised cases. The *repair rule* definition contains the repair rules. These are used to modify the solution retrieved from the case-base making it more suitable for the current situation. Both the *modification definition* and the *repair rule definition* may be omitted. To be a complete CBR system, however, it should contain both. The last set of blocks are the case instances. These make up the case base. The case file must contain at least one case instance and will initially need to be seeded with many cases before it is operable.

The following three stages are indicative of the case matching process (figure 5);

1. Weight rules may be applied to find a set of appropriate weights for performing case-matching.

2. The index values, which are either taken from the user case or specified separately by the user, are used to perform an index search. This retrieves a subset of cases from the case base which match all the index values exactly (except when abstraction symbols are specified as index values, in the *modification rules*).
3. Once this list of cases has been retrieved the user can allow the program to automatically select a case. This is based on weight-matching. For each case in the subset the case-matcher finds a weight which is obtained by totalling the weights of all matched fields. Fields which do not match exactly, but are defined to be similar by the modification rules, return a value which is less than the field's normal weight. The case-matcher selects the case with the highest total weight. The user can browse through the selected cases and select a case manually.

Figure 5. The Stages Involved in the Case Matching Process for ZONATION.



The method of case-matching in ZONATION consists of two phases. First, the enumeration-type fields cited in the index block are used to select a sub-set of cases from the case-base. Second, a form of nearest-neighbour (other types include interquartile distance, discrimination networks and parallel retrieval (Leake 1995)) matching is used to select the best case from the subset.

The weights are not attached to the cases themselves. ZONATION parses through each case in the subset evaluating their weight. A record of the best matching cases are recorded. The importance of each field is defined in the case definition section. ZONATION uses internal rules (not to be confused with the weight rules block) to evaluate what proportion of the weight is returned for each field. If for example, the values match exactly then the full weight is returned. In comparison, if two enumeration symbols are similar then 0.75 of the field weight is returned. Strings have to match exactly or zero field weight is returned. During the parsing, of two lists of symbols and for example, if half of them match, then half of the field weight is returned.

After the case has been selected extra values may be input if the case has a local field definition associated with it and then the global repair rules (in the *repair rule definition*) are enforced. Furthermore, if there are *local repair rules* associated with the case then these will be enforced. If a repair rule causes a reselection to occur, another case is selected using weight matching and local fields may again need to be entered. The repaired case is displayed and the user is given the option of adding the repaired case to the case base. If the user adds the repaired case to the case base, the values for the name and result of the case are entered by the user the case is then appended to the case file and added to the case base.

The following table is an example of a case definition for the ZONATION case file;

Table 1. A case definition for ZONATION

```
field land_element type is (foot_sunny, shoulder_sunny, foot_shady, shoulder_shady, rolling_rise,
rolling_hollow, bluff_sunny, bluff_shady) weight is 20;
field slope type is number weight is 15;  ~degrees
field aspect type is number weight is 15; ~degree_magnetic
field soil_depth type is  number weight is 12;   ~cm
field B_tex type is (loamy_sand, coarse_sandy_loam, sandy_loam, sandy_clay_loam, loamy_silt,
silt_loam, loamy_clay, missing_data) weight is 10;
field A_tex type is (loamy_sand, coarse_sandy_loam, sandy_loam, sandy_clay_loam, loamy_silt,
silt_loam, loamy_clay) weight is 0;
field A_horizon_depth type is  number weight is 0;   ~cm
field bulk_density type is number weight is 0;   ~g/cm
field dung_freq type is number weight is 0;   ~1to10
field volume_of_gravel type is  number weight is 0;
field weight_of_gravel type is  number weight is 0;
field weight_of_carbon type is  number weight is 0;
field volume_of_carbon type is  number weight is 0;
end;
```

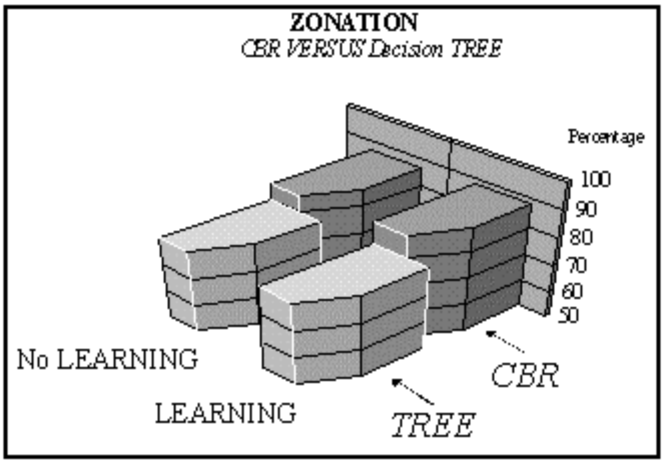
Spatial properties are defined as fields in the case file above. The case definition was used as a mechanism to process the spatial data input and the case instance was searched to fulfil the criteria of the case definition. Once a similar case instance was found it was possible to locate the similar case instance based on the fields of the case instance. Once these similar cases are located they can then be mapped and displayed. This shows the benefit a CBR provides, by increasing the GIS's analytical functionality and by adding an ability to learn. The traditional approach employed by soil scientists for soil classification uses an identification tree structure (Hewitt 1995). This example demonstrates the approach of using the GIS-CBR hybrid for soil classification.

A GIS-CBR hybrid has the ability to use explicit experiences to aid soil classification for new areas with similar spatial zones and attributes. These experiences (cases) enable the CBR to provide similar solutions (classification) to similar cases. Importantly, CBR differs from algorithms as no one solution is offered (as with an algorithm) and the user can choose a similar case from the solution set. Algorithms also need all criteria to be fulfilled while CBR allows some fields to be left blank without jeopardising the result. Results obtained through CBR improve after each new case is added to the case base. In comparison, an algorithm will predict the same answers and the associated error level each time it is used.

Explanation facilities are also easy to implement for a CBR system. The system will be used by those who feel comfortable with reasoning by analogy and who trust justifications that use data observations of past soil series incidents to support proposals, instead of chain rules that are triggered by *abstracted* threshold values. This explanation gives a system more chance of acceptance as ultimate responsibility for the decisions remains with the users. CBR systems can continuously incorporate new data in the form of cases and, in this way, adapt to long-term trends including soil degradation, loss and regeneration.

The success of the application was determined by comparing Hewitt's model with *ZONATION* against 200 observed plots. A summary of the results are presented in the following graph. Hewitt's model while predicting 200 plots produced a 20% error. *ZONATION* was run twice. During the first prediction correct adapted cases were not saved as new cases and an error of 13.5% was produced. During the second prediction correct adapted cases were saved as new cases and the error level was reduced to 12%. The errors were then scrutinised for patterns and it was found that soils with a clay content, and with a low B_horizon value, were difficult to classify using either Hewitt's or the ZONATION model. Evaluating these errors will allow scientists to further quantify automatic soil classification problems. These patterns aid in strengthening the decision tree classification which Hewitt used and in facilitating extra rules to be added to ZONATION. It was found, for example, that soils which had sandy loam clay produced more errors when trying to predict its soil type using the decision tree. It is, therefore important to create a new branch in the tree if the soil type is sand clay loam. Alternatively, within the case base the addition of a trigger would allow for deeper case matching to try and eliminate the cause of these errors.

Table 1. Performance Graph



Whilst noting that ZONATION had a 12% error rate and evaluating the comparison between the observed values and zonation the following points should be considered;

- The performance is indicative to the software used.
- The diverseness and harshness of the central Otago environment. Landslides are frequent, especially during torrential rain, because of poor soil and the gradient of the terrain. During land slides both the source area and the region where the soil was deposited are temporally changed. Thus the Irvin (1995) and Hewitt (1995) proposal that certain soils must nest in certain landforms does not hold. Over time it is likely that the soil should again nest in its land form due to physical geomorphic process.
- A degree of uncertainty is associated with the initial control observations due user and instrument errors.
- Currently, this model is being applied to tourist spatial movements and the case history of aeroplane accidents (Higham *et al.* 1996).

CONCLUSION

Environmental problems are inherently complex. This research has proposed a novel method (GIS-CBR hybrid) to aid in modelling and solving of such problems. This CBR-GIS hybrid benefits from the functionality of both systems. Selecting spatial cases using GIS functionality (proximity, connectivity, adjacency) are such examples.

Finding similar spatial cases with certain fields and in a certain proximity to or adjacency, connectivity to a spatial phenomena using CBR functionality. A case-based approach is beneficial as CBR systems have the ability to continually learn and evolve through the capturing and retainment of past experiences. This paper also illustrates the potential of AI in the spatial realm in recognising sets of patterns, predicting, providing decision support and simulating spatial phenomena. This includes recognising situations and structuring data to give spatial solutions to spatial problems. It is suggested that GIS would benefit from a CBR link. This is the first stage of *ZONATION* and an attempt to display the concept of CBR for soil modelling. It also attempts to indicate the possible extension of this concept to other environmental concerns such as geology, natural hazards and vegetation cover. The next stage of *ZONATION* should be more spatially oriented and incorporate graphical user input.

REFERENCES

Aamodt, A. and E. Plaza, 1994. Case-based Reasoning: Foundational Issues, Methodological Variations and System Approaches. *Artificial Intelligence Communications,* Vol.7 No.1.

Aha, D. W. 1994. Case-Based Reasoning, AIRIES 94, Workshop. Biloxi, MS.

Althoff, K. Wess, S. and M. Manago, 1994. Reasoning With Cases-Theory and Practice ECAI.

Benwell, G. L. 1996. A Land Speed Record? Some Management Issues to Address, The Tewkesbury Lecture, and Invited Keynote Address, *Proceedings of the Conference on Managing Geographic Information For Success*, Conference Proc, Department of Geomatics, The University of Melbourne, pp70-75.

Berger, J. 1994. Roentgen: Radiation Therapy and Case-based Reasoning. *In the Proceedings of the Conference on Artificial Intelligence Applications,* Proceedings. pp.171-177.

Branting, K. L. and J. D. Hastings, 1994. An Empirical Evaluation of Model-based Case Matching and Adaption. *American association for artificial intelligence, case-based reasoning workshop,* Proceedings. Seattle, Washington.

Burrough, P. A. and A. U. Frank, 1995. Concepts and Paradigms In Spatial Information: are Current Geographic Information Systems Truely Generic? Vol.9. No.2. pp.101-116.

Burstein, F. and H. Smith, 1994. Case-based Decision Making for Intelligent Decision Support. Department of Information Systems, University of Monash.

Fischer, M. M. and P. Nijkamp, 1993. Design and Use of Geographic Information Systems and Spatial Models. In: *Geographical information systems, spatial modelling, and policy evaluation*, Eds. M. M. Fischer and N. P. Springer Verlag, Berlin.

Frank, A. U. 1996. Qualitative Spatial Reasoning: Cardinal Directions as an Example *International Journal of Geographical Information Systems,* Vol.10. No.3. pp.269-290.

Hernandez, D. 1993. Reasoning with Qualitative Representations: Exploiting the Structure of Space. In Piera Carrete and Singh pages 493-502.

Hewitt, A. E. 1995. Soils of the Conroy Land System, Central Otago. Lincoln, Canterbury, New Zealand, Manaaki Whenua Press. Landcare Research Science Series.

Higham, E.C., Holt A. & G.W. Kearsley 1996. Tourist Flow Reasoning: The Spatial Similarities of Tourist Movements. In the Proceedings of the 8th Annual Colloquim of the Spatial Information Research Centre, Otago University, Dunedin, New Zealand, pp 69-78.

Holt, A. Benwell, G. L. & Kasabov, N. K. (In press) The Case-Based Reasoning (CBR) Approach, *Hand Book for Geostatistics and Spatial Data Analysis: Elements of Interpretation and Examples*, Editors Shibli, S. A. R. Michel, M. & D. Gregoire.

Holt, A. and G. L. Benwell, 1996. Case-based Reasoning and Spatial Analysis. *Journal of the Urban and Regional Information Systems Association,* Vol.8. No.1. pp 27-36.

Holt, A. 1996. Allowing the Environment to Model Itself. Environmental Perspectives, A Triannuual Newsletter published by the Environmental Policy and Management Reseach Centre. Issue 10.

Holt, A. and G. L. Benwell, 1995a. Spatial Reckoning using Case-based Reasoning. *8th Australian Artificial Intelligence conference incorporating the Workshop on AI in the environment,* Proceedings. Canberra, pp. 53-66.

Holt, A. and G. L. Benwell, 1995b. Case-based Reasoning with Spatial Data. *The 2nd New Zealand International two-stream conference on Artificial Neural Networks and Expert Systems,* Proceedings. Eds. N.K. Kasabov and G. Coghill. Dunedin, New Zealand, IEEE Computer Society Press, pp. 385-388.

Irvin, B. J. Ventura, S. J. & B.K. Slater, 1995. Landform Classification for Soil-landscape Studies. *Annual ESRI User Conference.*

Jankowski, P. 1995. Integrating Geographical Information Systems and Multiple Criteria Decision-making Methods. Vol.9. No.3. pp.251-273.

Jones, E. K. and A. Roydhouse, 1994. Spatial Representations of Meteorological Data for Intelligent Retrieval. *The Sixth Annual Colloquium of the Spatial Research Centre,* Proceedings. Eds. G.L. Benwell and N.C. Sutherland. Dunedin, New Zealand. pp.45-58.

Kasabov, N. and R. L. Trifonov, 1993. Using hybrid Connectionist Systems for Spatial Information Processing. *The Fifth Annual Colloquim of the Spatial Research Centre,* Proceedings. Eds. Benwell G. L. and N. C. Sutherland. Dunedin, New Zealand.pp85-96.

Kasabov, N. K. and D. Nikovski, 1992. Prognostic Expert Systems on a Hybrid Connectionist Environment. In: *Artificial intelligence- methodology, systems, applications*, Eds. B. Du Boulay and A. V. S. Gurev. Elsevier, Holland.

Kolodner, J. 1993. Case-Based Reasoning. San Mateo, Morgan Kaufmann Publishers.

Laurini, R. and D. Thompson, 1992. Spatial Knowledge. In: *Fundamentals of Spatial Information Systems*, Academic Press, London.

Leake D. B. 1995. Case-Base Reasoning: Issues, Methods and Technology, *A Tutorial for the First International Conference on Case-Based Reasoning*, Sesimbra, Portugal.

Lekkas, G. P. N. M. Avouris, & L.G. Viras, 1994. Case-Based Reasoning in Environmental Monitoring Applications. *Applied Artificial Intelligence An International Journal,* Vol.8. No.3. pp.359-376.

Leung, Y. and K. S. Leung, 1993. An Intelligent Expert System Shell for Knowledge-based Geographical Information Systems: 1. The Tools. *International Journal of Geographical Information Systems,* Vol.7. No.3. pp.189-199.

Malek, M. and A. Labbi, 1995. Integration of Case-based Reasoning and Neural Networks Approaches for Classification.

Openshaw, S. 1995. Human Systems Modelling as a New Grand Challenge Area in Science (Commentary). *Environment and Planning A,* Vol.27. pp.159-164.

Openshaw, S. 1993. Modelling Spatial Interaction Using a Neural Net. In: *Geographical information systems, spatial modelling, and policy evaluation*, Eds. M. M. Fischer and N. P. Springer Verlag, Berlin.

Skidmore, A. 1995. Neural Networks and GIS. *GIS User,* No.11 (May-July, 1995), pp.53-55.

Skidmore, A. K. Baang, J.S. and P. Luckananurug, 1992. Knowledge-based Methods in Remote Sensing and GIS. *The 6th Australasian Remote Sensing Conference,* Proceedings. Wellington, New Zealand, pp. 394-403.

Skidmore, A. K. Ryan, P. J. Dawes, W. Short, D. and E. O'Loughlin, 1991. Use of An Expert System to Map Forest Soils from a Geographical Information System. *International Journal of Geographical Information Systems,* Vol.5. No.4. pp.431-445.

Smith, T. R. and Yiang, J.E. 1991. Knowledge-based Approaches in GIS. In: *Geographic Information Systems*, Eds. D.J. Maguire, M. F. Goodchild and D. Rhind. Longman Scientific and Technical, Essex.

Webster, C. 1990. Rule-based Spatial Search. Vol.4. No.3. pp.241-259.

Williams, G. J. 1995. Templates for Spatial Reasoning in Responsive Geographic Information Systems. *International Journal of Geographical Information Systems,* Vol.9. No.2. pp.117-131.