

# SES: A Soils Expert System.

A Thesis presented  
in fulfilment of the requirements  
for the degree of  
Master of Science  
in  
Computer Science

by

E. G. Todd

July 1988

Massey University

### **Abstract**

An Expert System Development Methodology is proposed, based on experimentation in developing a soils expert system (SES) which identifies a soil from incomplete field data. Tools for conceptual modeling of the soils domain are examined. The tools developed provide a means of recording the conceptual model of the knowledge from three different view points: the inference structures, the domain objects and the functional aspects. A review of the structures used in the knowledge bases of existing classification problems identifies eleven categories for grouping these structures. Using this information with the conceptual model, a detailed design of the knowledge base for SES is created. This design closely models those structures identified as important in the soils domain ensuring that important knowledge is represented explicitly.

## Acknowledgements

This thesis is the result of a research program in which the author is indebted to the labours of many people and organizations. I would like to express my appreciation to Mr Ray. Kemp, my Supervisor, for his guidance and motivation throughout the time of this research. I wish to especially thank my husband, Arthur, for continually supporting and encouraging me to complete this piece of research.

I would also like to thank :

Dr Alan Palmer (Pedologist, Department of Soil Science, Massey University) who acted as the main domain expert and also Dr Bob Stewart (Geologist, Department of Soil Science, Massey University) and Mr Hugh Wilde (Pedologist, DSIR Soil Bureau) who both made contributions to the knowledge base of SES.

Members of the Staff of the Computer Science Department, who have given advice, assistance and encouragement.

## Table of Contents

Abstract.....	ii
Acknowledgements .....	iii
Table of Contents.....	iv
List of Figures.....	viii
 Chapter 1	
Introduction.....	14
1.1 Purpose of the Project .....	16
1.2 The Soil Science Domain.....	16
1.3 Organization of the Study. ....	19
 Chapter 2	
An Overview of Expert Systems.....	20
2.1 What is an Expert System .....	20
2.2 Structure of an Expert System.....	23
2.2.1 The knowledge Base.....	24
2.2.2 The inference engine.....	26
2.2.2.1 The control mechanism.....	26
2.2.2.2 The inference mechanism.....	27
2.2.3 The user interface.....	29
2.2.4 The explanation module.....	29
2.2.5 Knowledge Acquisition Facility .....	30
2.3 Knowledge Representation .....	31
2.3.1 Rules .....	33
2.3.2 Objects .....	36
2.3.3 Logic .....	37
2.4 Developing an Expert System.....	38
2.4.1 Prototyping.....	38
2.4.2 Stages in Prototype Development.....	39
2.4.3 Prototyping and Expert System Development.....	41
2.5 Summary .....	43

## Chapter 3

The Feasibility of SES.....	44
3.1 Deciding on the Feasibility of a Project.....	45
3.1.1 A Soils Identification Expert System.....	47
3.2 Development of the Initial Prototype.....	48
3.3 Developing the Second Prototype.....	50
3.2.1 The Interface.....	50
3.3.2 Uncertainty.....	53
3.3.3 Assessment.....	54
3.4 Criteria for further development of SES .....	54
3.5 Summary .....	55

## Chapter 4

The Conceptualization Stage.....	57
4.1 Steps in the Conceptualization Stage .....	57
4.2 Specifying Requirements from the User's Perspective .....	58
4.2.1 The Domain Expert's Requirements .....	59
4.2.2 The Student User's Requirements.....	60
4.3 Specifying the Knowledge Requirements of the Modules.....	61
4.3.1 Inference Module.....	61
4.3.2 Explanation Module.....	62
4.3.3 Communications Module.....	65
4.3.4 Acquisition Module .....	70
4.4 Summary .....	71

## Chapter 5

The Conceptual Model of SES.....	72
5.1 Specification of an Expert System .....	72
5.1.1 High Level Specification using Inference Structures .....	75
5.1.2 Specifying Domain Objects.....	77
5.1.3 Specifying the Functional Aspects.....	80
5.2 Developing a Conceptual Model.....	81
5.3 The High-level Specification of SES .....	82
5.3.1 Specific Inference Paths in the Soils Domain.....	83
5.3.2 A Generalized Inference Structure for SES.....	84
5.4 The Domain Objects.....	86
5.4.1 The Generalization Plane.....	87
5.4.2 The Aggregation Plane .....	89
5.4.3 Informal Classifications used in the Soils Domain.....	93

5.5 Functional Aspects.....	94
5.6 Summary .....	96
Chapter 6	
The Formalization Stage .....	98
6.1 Knowledge Structures in MYCIN.....	99
6.2 Knowledge Structures in PROSPECTOR.....	102
6.3 Knowledge Structures in CENTAUR .....	106
6.4 Summary .....	110
Chapter 7	
The Design Phase of SES.....	114
7.1 Using Prototyping to Experiment with Different Knowledge Structures for Representing the Soils Domain.....	114
7.1.1 Redesigning the structure of the Knowledge Base .....	116
7.1.2 Evaluating the Changes.....	118
7.1.3 Expanding the Prototyping .....	119
7.2 Structural Design for SES .....	123
7.3 Specifying the Knowledge Structures for SES .....	126
7.3.1 The Knowledge-Structure Schema .....	128
7.3.2 The State Schema.....	130
7.3.3 The Classification Schema.....	132
7.3.4 The Soil Description Schema.....	135
7.3.5 The Type Object Schema.....	137
7.3.6 The Auxiliary Schema .....	138
7.4 Summary .....	140
Chapter 8	
Summary and Conclusions.....	142
8.1 Expert System Development Methodologies.....	142
8.2 The role of Prototyping in Expert System's Development.....	143
8.3 Specifying the Knowledge Base. ....	144
8.4 Expert Systems and Soil Science .....	146
8.5 Directions for further Research.....	147
8.6 Conclusion.....	148
Appendix A	
Examples from Prototyping Exercise 1.....	149
A.1 Example Dialogue of Consultation Session.....	149
A.2 The Dynamic Database. ....	150
A.3 The Knowledge base. ....	151

Appendix B	
Examples from the Menu-Driven Prototype.....	154
B.1 Example of a voluntary Data entry Session.....	154
B.2 Examples of the Help Screens.....	161
Appendix C	
Selected Sources for Soils Knowledge.....	162
Appendix D	
Design Aspects of the Generalization Plane for SES.....	169
Appendix E	
Design Aspects of the Aggregation Plane for SES. ....	173
Appendix F	
Functional Design Aspects of SES.....	178
References .....	186

## List of Figures

Figure 1.01	Hierarchy of Soil Classification Units. ....	18
Figure 2.01	The four primary modules of an Consultation System showing the main communication channels.....	24
Figure 2.02	Comparison of Traditional Life Cycle with.....	42
Figure 3.01	Generic categories of knowledge engineering applications.....	45
Figure 3.02	Generic operations for synthesizing and analyzing a system.....	46
Figure 3.03	A Type 1 Menu .....	51
Figure 3.04	A Type 2 Menu .....	51
Figure 3.05	A Yes/No Type 1 Question.....	52
Figure 3.06	A Type 2 Question .....	52
Figure 3.07	Rule showing premises with unequal weights. ....	53
Figure 4.01	Rules from the Explanation Prototype.with the associated Inference chain.....	63
Figure 4.02	Examples of dialogue boxes.....	67
Figure 4.03.	Block diagram of Otaki district showing the Tokomaru Terrace system highlighted .....	68
Figure 4.04	Diagram for soil texture assessment. ....	69
Figure 4.05	Percentage diagrams.....	69
Figure 4.06	Summary of the kinds of knowledge required by the different modules.....	71
Figure 5.01	The RUN-DEBUG-EDIT cycle related to the ISF. ....	74
Figure 5.02	Inference structure of heuristic classification. ....	76
Figure 5.03	A Semantic Network Diagram.....	79
Figure 5.04	A State Specification Diagram.....	81
Figure 5.05	Generalizations from specific inferences.....	83
Figure 5.06	Specific inference diagram.....	84
Figure 5.07	Levels of heuristics matching in the soils domain. ....	85
Figure 5.08	Identification of the main Systems from the generalization of inference paths. ....	85
Figure 5.09	Generalized inference structure of knowledge about Soil identification. ....	86
Figure 5.10	Hierarchy for soils data.....	86



Figure 5.11	Semantic Net of the Schema for the Generalization Plane. ....	88
Figure 5.12	Specific Semantic Net showing part of the Soils Classification Network.....	89
Figure 5.13	Semantic Net of the Schema for the Aggregation Plane.....	90
Figure 5.14	Key to symbols used to designate specific relationships used in the aggregation plane diagrams. ....	90
Figure 5.15	Semantic Net of the Schema for the Horizon Descriptor.....	91
Figure 5.16	Semantic Nets showing the schema and a specific instance of the object describing the Structure of an Aggregate.....	92
Figure 5.17	Semantic net showing two of the partitions the node "Otaki terrace" participates in. ....	93
Figure 5.18	Semantic net with partitions for the special association relationship.....	94
Figure 5.19	Initial Simplified State Specification Diagram for SES.....	95
Figure 5.20	Simplified State specification diagram for SES.....	95
Figure 6.01	Object-Attribute-Value triples .....	99
Figure 6.02	Mycin like rule showing functions and O-A-V triplets with English like translation. ....	100
Figure 6.03	Context Template.....	101
Figure 6.04	Parameter Template. ....	101
Figure 6.05	Semantic network showing the representation of the "overlies" relationship.....	103
Figure 6.06	Network showing subset relationships based on a physiographic division of soils. ....	104
Figure 6.07	Net showing that composition is a relation that has an instance linking parent material to volcanic tuff.....	104
Figure 6.08	The nodes within the box form the partition for the delineation of the composition relations. ....	105
Figure 6.09	Network for a rule. ....	105
Figure 6.10	Network of Prototype frames based on generic classification.....	107
Figure 6.12	Component frame defining physiographic features characteristic of a Yellow Grey Earth.....	109
Figure 6.13	Kinds of knowledge from the three knowledge bases - MYCIN, PROSPECTOR, CENTAUR. ....	113
Figure 7.01	Two of the rules created from one parent rule in the second prototype .....	115
Figure 7.02	Two rules with a subset relationship.....	115

Figure 7.03	Rule using general comparisons of soil properties. ....	116
Figure 7.04	Part of the Semantic Net for a Yellow Grey Earth with modified tuple representation for A-Horizon. ....	117
Figure 7.05	Soil groups known to the main prototype. ....	120
Figure 7.06	Ranking of Soil Group Classification, ....	121
Figure 7.07	Tuple example showing sufficiency and necessity. ....	122
Figure 7.08	Shows groupings of different types of knowledge in the Knowledge Base. ....	123
Figure 7.09	Part of the Knowledge Structure Hierarchy showing links to specific instances. ....	128
Figure 7.10	Frame describing the Root Schema of the Knowledge Structure Hierarchy.....	129
Figure 7.11	Section describing the Instance structure of the State Schema.....	131
Figure 7.12	Part of the frame describing the Start-state.....	132
Figure 7.13	Link section of the frame describing one classification unit in the generalization plane. ....	133
Figure 7.14	Part of the frame describing the Shannon Series. ....	134
Figure 7.15	Link section of the frame describing the schema for one descriptor in the aggregation plane. ....	135
Figure 7.16	Defining the Features section in the instance-structure section of the site descriptor schema. ....	136
Figure 7.17	The Features section Shannon series site descriptor.....	136
Figure 7.18	Part of the frame describing the type-object for drainage.....	138
Figure 7.19	The two types of tuples used to create semantic net for field data. ....	139
Figure 7.20	Part of a semantic net recording field data.....	139
Figure 7.21	Kinds of knowledge required for SES. ....	141
Figure C.01	Typical Profile description. ....	162
Figure C.02	Main soil profile features.....	163
Figure C.03	Page 1 of soil profile recording card for recording soil forming factors.....	164
Figure C.04	Page 2 of soil profile recording card for recording profile characteristics.....	165
Figure C.05	Part of a soil map legend. ....	166
Figure C.06	Diagram showing the relationships between climate, topography and soil development. ....	167
Figure C.07	Diagram showing some of the soil forming processes.....	168

Figure D.01	Part of the Knowledge Structure Hierarchy with classification subtree highlighted.....	169
Figure D.02	Part of the Semantic Net for the Schema for the Generalization Plane.....	169
Figure D.03	Frame representing the classification schema which describes the common structures of the object-schemas for each classification unit. ....	170
Figure D.04	Example of a classification object-schema. This frame is the Series Schema which describes the structure of the frames representing each of the Soil Series known to SES.....	171
Figure D.05	Example of an instance of a Series Schema. This frame describes the Shannon Series of soils. ....	172
Figure E.01	Part of the Knowledge Structure Hierarchy with soil description subtree highlighted.....	173
Figure E.02	Semantic Net of the Schema for part of the Aggregation Plane.....	173
Figure E.03	Frame representing the soil description schema which describes the common structures of the object-schemas for each object in a soil descriptor.....	174
Figure E.04	Example of an instance of a site descriptor. This frame describes the general features required to characterize one of the soil Series.....	175
Figure E.05	Example of an instance of a site descriptor. This frame describes the site of a Shannon Series soil. ....	176
Figure E.06	Frame representing the Type-object Schema which describes the common structures of the object-schemas for each type required by the SES.....	176
Figure E.07	Example of an instance of a Type-object Schema. This frame describes the soil feature, drainage.....	177
Figure F.01	Part of the Knowledge Structure Hierarchy with part of the state specification subtree highlighted. ....	178
Figure F.02	State Specification diagram for SES.....	179
Figure F.03	Frame describing the Root Schema of the State Schema Hierarchy.....	180
Figure F.04	Frame describing the Start State for a consultation in SES. ....	181
Figure F.05	Frame describing the state for testing whether the essential features of the candidate solutions match those of the field data. ....	182
Figure F.06	Frame describing the state for checking that the profile characteristics of the candidate solutions match those of the field data. ....	183

Figure F.07	Frame describing the state for checking associated and similar solutions which may better match the field data.....	184
Figure F.08	Frame describing the state for checking frames lower in the solution heirarchy.....	184
Figure F.09	Frame for describing conditions for asking the user if they would like advise. ....	185
Figure C.02	Main soil profile features.....	
Figure C.03	Page 1 of soil profile recording card for recording soil forming factors. ....	
Figure C.04	Page 2 of soil profile recording card for recording profile characteristics. ....	
Figure C.05	Part of a soil map legend. ....	
Figure C.06	Diagram showing the relationships between climate, topography and soil development.....	
Figure C.06	Diagram showing some of the soil forming processes.....	
Figure D.01	Part of the Knowledge Structure Hierarchy with classification subtree highlighted.....	179
Figure D.02	Part of the Semantic Net for the Schema for the Generalization Plane.....	179
Figure D.03	Frame representing the classification schema which describes the common structures of the object-schemas for each classification unit. ....	180
Figure D.04	Example of a classification object-schema. This frame is the Series Schema which describes the structure of the frames representing each of the Soil Series known to SES.....	181
Figure D.05	Example of an instance of a Series Schema. This frame describes the Shannon Series of soils. ....	182
Figure E.01	Part of the Knowledge Structure Hierarchy with soil description subtree highlighted.....	183
Figure E.02	Semantic Net of the Schema for part of the Aggregation Plane.....	183
Figure E.03	Frame representing the soil description schema which describes the common structures of the object-schemas for each object in a soil descriptor.....	184
Figure E.04	Example of an instance of a site descriptor. This frame describes the general features required to characterize one of the soil Series.....	185

Figure E.05	Example of an instance of a site descriptor. This frame describes the site of a Shannon Series soil. ....	186
Figure E.06	Frame representing the Type-object Schema which describes the common structures of the object-schemas for each type required by the SES.....	187
Figure E.07	Example of an instance of a Type-object Schema. This frame describes the soil feature, drainage.....	189
Figure F.01	Part of the Knowledge Structure Hierarchy with part of the state specification subtree highlighted. ....	189
Figure F.02	State Specification diagram for SES.....	190
Figure F.03	Frame describing the Root Schema of the State Schema Hierarchy.....	191
Figure F.04	Frame describing the Start State for a consultation in SES. ....	192
Figure F.05	Frame describing the state for testing whether the essential features of the candidate solutions match those of the field data. ....	193
Figure F.06	Frame describing the state for checking that the profile characteristics of the candidate solutions match those of the field data.....	194
Figure F.07	Frame describing the state for checking associated and similar solutions which may better match the field data.....	194
Figure F.08	Frame describing the state for checking frames lower in the solution heirarchy.....	195
Figure F.09	Frame for describing conditions for asking the user if they would like advise. ....	195

## Chapter 1: Introduction

During the nineteen sixties and early nineteen seventies Artificial Intelligence research concentrated on the development of general purpose problem solving strategies. The best known example is the General Problem Solver (GPS) built by Newell, Shaw and Simon (1963). It attempted to replicate the kind of problem solving humans use every day. From the nineteen seventies the direction of this research has altered. This change was initiated by the publication of research that concentrated on the development of programs for solving particular problems in specialized areas of domain knowledge. Such programs are currently referred to as expert systems.

Specifically, the research related to the MYCIN expert system (Shortliffe, 1976), triggered interest in the application of the results of Artificial Intelligence research to the wider software development arena. In the early 1980's the products of the Artificial Intelligence research centres have been taken up by innovative development units in both universities and industry. The main direction of this research has been towards the development of effective expert systems.

Medical domains are the most widely represented in the expert system literature. MYCIN is the best documented of these systems, and contains knowledge about the family of meningitis diseases. It has been extensively tested and modified. Research following on from MYCIN has led to a whole family of expert systems and expert system shells (Clancey,1986). EMYCIN (Empty MYCIN), is one of the earliest expert system shells. EMYCIN was built by removing the knowledge from the knowledge base of MYCIN. PUFF (Kunz et al, 1978) was built using EMYCIN by adding a knowledge base about pulmonary physiology. CENTAUR (Aikins,1983) was built by taking the knowledge in PUFF and redesigning the knowledge structures to represent the same knowledge using a mixture of frames and rules

Expert system technology has also been used successfully by the computing industry itself. The most notable example is R1 (McDermott, 1980), which helps to configure Digital Equipment Corporation's VAX series of computers. DIGITAL has claimed that this program has allowed it to gain a significant advantage over competitors. Other computer companies have followed this lead. PRIME has developed an expert system DOC (Littleford,1985) which analyses a memory image from a crashed Prime computer system. The expert system deduces the cause of the crash and recommends a course of action for repair. DOC only loads that part of the knowledge base that is relevant to the

problem after ascertaining what model of CPU is involved and the significant peripheral devices.

The application of expert system technology is only just beginning to be widely reported. A literature search reveals that expert system technology is being applied in the space industry, in business management, the oil industry and by the military. An important application of expert system technology is in interpreting and applying complex codes of practice and specific sets of regulations. In New Zealand, BRANZ (Whitney, 1987) has successfully developed a system to help check that a building design complies with the fire regulations. Government Computing Services (Barton, 1987) has developed a system to help determine a client's unemployment benefit eligibility.

Much expert systems based research has been into knowledge representation. This has been developed in parallel with natural language processing. Both these lines of research are important for the development of large computerized data stores. Current data bases are limited to factual knowledge and lack the semantic and heuristic knowledge of an expert system knowledge base. Although special purpose query languages have been developed for accessing the data, often potential users are either unsure of exactly what they are searching for, or alternately, how to phrase the questions so the required information can be obtained. Natural language research has enabled the development of natural language interfaces to a number of database products. These help users access the information they require. Examples of a number of systems are outlined in Bonnet (1985).

The retrieval of computer based information and the use of application packages by non-computing personnel provides a diverse area for the application of expert systems technology in the commercial environment. Expert systems can be built for existing computerized knowledge sources and application packages. For these expert systems the domain of expertise is a combination of knowledge about the application and about how people typically wish to use the application. SACON (Bennett and Englmore, 1979) is a front end to an application which determines the resistance of different materials. SACON helps the user by giving advice on how to use the system for analysing structures.

Expert system technology is extending the limits to the type of knowledge that can be stored on the computer and the way in which this knowledge can be used. Expert system technology has expanded users expectations concerning the type of information they can request from a computerized system. In particular expert systems have knowledge of how they work and can therefore give some explanation of their actions and lines of reasoning.

### 1.1 Purpose of the Project

The purpose of this project is to investigate the suitability of using expert system technology in the field of soil science. The main aim is to evaluate relevant methodologies for analysis and design of expert systems. This study emphasizes the role of prototyping in expert systems development and the design of the knowledge base. Development of diagrammatic tools for developing and representing models of different aspects of the knowledge base form a significant part of this research.

The target application area is the identification of New Zealand Soils from incomplete field data. Not only should the system be able to identify a soil accurately from sufficient data but also it should be able to report when it is not possible to determine an unambiguous identification. At this point the system should offer help to the user by identifying the additional data that is required. It should highlight the important features that would either confirm or negate the most likely candidate soil types provisionally identified from the data so far.

### 1.2 The Soil Science Domain

The areas to which expert system technology have been applied are expanding rapidly. Systems that are used in specialized aspects of the earth sciences were one of the earliest application areas. PROSPECTOR, (Duda, Gaschnig and Hart, 1979) evaluates geological structures for the purpose of identifying and assessing the commercial viability of mineral deposits. The inputs to the system are the geological field data collected by geologists and the output is a site evaluation and maps of the deposit.

Other earth science associated expert systems include: DRILLING ADVISOR (Harmon and King, 1986), DIPMETER ADVISOR (Baker, 1984) and LITHO (Ganascia, 1984). DRILLING ADVISOR provides advice on solving problems encountered with drill bits when drilling exploration and production drill holes. The other two systems are used to interpret the data from down hole wire-line logging of drill holes, particularly in petroleum prospecting.

Massey University has specific expertise in pedology<sup>1</sup>, one of the branches of the earth sciences. With the diversification of agriculture and horticulture in New Zealand the application of computer technology to the dissemination of information about soils is a timely project to tackle.

---

<sup>1</sup> Pedology is the study of Soil Science.



The identification of a soil from field data has a number of parallels with the evaluation of geological field data. Both disciplines are basically concerned with describing three-dimensional layers including, their characteristics, boundary conditions and the processes involved in their formation. Both disciplines generalize the descriptions by producing classification systems; systems which permit specific instances of a layer or group of layers to be sorted, compared, correlated and contrasted.

Soils, like rock formations, are the product of the intersection of a number of closely interrelated processes. These processes are not discrete but progressively change over time and in space. Creating an hierarchical classification of soils is therefore inherently subjective. Classifying a specific soil also involves a degree of incidental association, as indicated by the following quote from Taylor and Pohlen (1970)

" a full definition of a kind of soil includes a statement of both differentiating and accessory characteristics, of the permissible ranges in each, and of any likely accidental characteristics that may serve as phase distinctions. "

Before a soil can be classified the pedologist has to describe the soil. A soil description records both the soil forming factors and the soil morphology. Soil forming factors have been recorded in the site descriptor since Dokuchaiev (Neill, Palmer and Pollock, 1987) observed that

" soils are products of extremely complex interactions of local climates, plants and animals, parent rocks, topography and the ages of landscapes"

The modern pedologist views the soil he can describe in the vertical section profile as a complete integrated, natural body that reflects the combined effects of the soil forming factors. From these direct observations pedologists have noted the associations between the site descriptors and the soil morphology. The association between the soil forming factors can be shown in a simplified soil-function equation

$$s = f(\text{cl, o, r, p, t})$$

where:

- s = soil
- cl = climate
- o = organisms
- r = relief
- p = parent material
- t = time

This equation is a surface reflection of the processes that form a soil. Many of the soil forming processes are as yet poorly understood and the soil forming factors associated with a process are often uncertain.

For example a New Zealand podzol (Neall, Palmer and Pollock, 1987) forms under the following conditions:

- cl - adequate rainfall in humid and superhumid regions
- o - under kauri forest in Auckland and to a lesser extent under rimu or beech trees
- r - flat or rolling relief, not on steep sites
- p - coarse or medium textures parent material
- t - and enough time for the expression of the process on the soil morphology

The site description holds information about these soil forming factors and the profile description holds the information about the soil morphology. A profile description is a detailed inventory of the changes in the major soil characteristics, beginning at the ground surface and extending vertically down to the underlying rock material. Each soil is made up of layers that are termed horizons. For a specific soil at the site where the profile hole is dug characteristics are recorded and indexed by depth via the horizon designation. For each horizon characteristics such as colour, consistency, porosity, size and shape of aggregates and degree of compaction are recorded to build up a detailed description of the form and structure of the profile.

The classification system used for soil identification purposes is derived from the soil mapping units used on the soil maps accompanying the DSIR Soil Bureau Bulletins. This classification system is a simple hierarchy that forms a pyramid of units, figure 1.1. The soil type forms the smallest unit while the largest unit is the soil class.

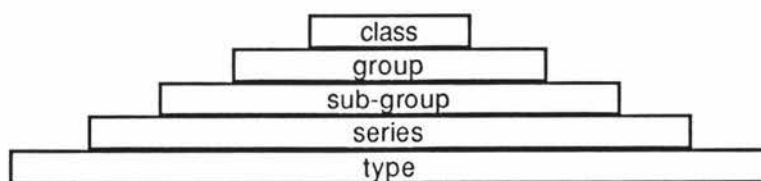


Figure 1.01 Hierarchy of Soil Classification Units.

The Soil Bureau bulletins are published on a County basis thereby forming spatially discrete units of knowledge. The information in these Soil Bureau bulletins and the associated maps is not easy to interpret correctly by either students of soil science or agriculture specialists. An expert system could complement this information by helping the user to interpret the field data they have collected. By identifying the most probable

classification unit, the system could enabling the user to make more effective use of detailed soil properties and characteristics in the bulletins.

### 1.3 Organization of the Study.

Chapter two reviews expert system technology. The review summarizes how expert systems can be classified. The different components of an expert system, of which the knowledge base is the central component, are described. The chapter concludes with a discussion of current expert system development methodologies.

The five stage methodology proposed by Hayes-Roth et al (1983) was selected to guide the development of the Soils Expert System, SES. Chapter three discusses Identification, the first of these stages. An assessment of the feasibility of the proposed expert system. Guide-lines for determining whether a problem domain is suitable for an expert system solution are reviewed. The use of prototyping to determine the feasibility of an expert system is explored.

The conceptualization stage defines both the requirements for the system and the specification of the conceptual design of the knowledge base. The requirements of the users and the modules comprising the expert system are outlined in chapter four. Prototyping is used to clarify specific problem areas, for instance interface design.

Aspects of the specification of a conceptual model for an expert system are reviewed in chapter five. Three views of the domain knowledge are identified as important for describing the conceptual model of a knowledge base. These are used to develop the conceptual model of the knowledge base of SES.

The structural or low level design of an expert system corresponds to Hayes-Roth's formalization stage. A review of the knowledge base structures used in three existing expert systems forms the basis of chapter six. The common features of these knowledge base structures are identified.

Information gained from identifying the common structures in existing knowledge bases was used to redesign the prototype for SES. The evaluation of this prototyping exercise is described in chapter seven. The description of the types of knowledge base structures considered necessary for SES are based on the results of this evaluation.

The final chapter contains a summary of the main points discussed in the thesis with reference to the stated aims of the research. Additional applications for expert system technology in the soil science field are suggested. Extensions and improvements to the current project and associated topics are outlined as suggestions for further research.