ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

SCHOOL OF INFORMATION STUDIES FOR AFRICA

APPLICATION OF EXPERT SYSTEMS IN SPECIES SELECTION: THE CASE OF FORESTRY RESEARCH CENTER (FRC).

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE IN INFORMATION SCIENCE

BY
SAMIR ABDUSELAM IBRAHIM
July 2001.

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

SCHOOL OF INFORMATION STUDIES FOR AFRICA

APPLICATIONS OF EXPERT SYSTEMS IN SPECIES SELECTION: THE CASE OF FORESTRY RESEARCH CENTER. (FRC)

BY

SAMIR ABDUSELAM

Name and Signature of Members of the Examining Board

Ato Getachew Jemaneh, Chairman, Examining Board ______________

Ato Tesfaye Biru, Advisor ______________

Ato Million Meshesha, Advisor ______________

Dr. Kamal Bechkoum, External Examiner ______________

DEDICATION

For my late parents

Abduselam Ibrahim & Fatuma Shash
" May Allah shower his mercy upon you both."

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

ACKNOWLEDGEMENT

Firstly, I praise ALLAH, the almighty, for his guidance. I also thank him for making those seemingly impossible things happen to me.

I thank my advisors Ato Tesfaye Birru and Ato Million Meshesha for their commitment, patience and comments to the progress and completion of my thesis work. I also want to acknowledge with thanks, the support of all SISA community for being kind and helpful – namely: - lecturers, Ato Getachew, Ato Dereje, Ato Workshet, Ato Ermias, Dr. Nega and W/ro Woinshet; and all my classmates, not forgetting Suleyman, Nadia and Tesfaye.

My Special thanks also goes to, Ato Mebratu Mihretu, Ato Dechasa Jirru and Ato Mesekere Reta, the experts in FRC for their motivation during series of interviews. I would also like to thank Ahmed Bedru, FRC librarian, Melaku and Henok, for providing me invaluable documents.

I would also like to thank my beloved family, especially my sisters, Meftuha, Gizman and Hamdia, my brothers, Testi, Dinir and Mohammed (Mahir), and lastly, auntie Aziza, uncle Kamil and their family for their love, appreciation and prayers for my achievement. I also express my earnest gratitude to my uncles- Abdullahi and his son Sami, Abdurahman and Abubeker, to their moral and financial support throughout the course of my study.

ABSTRACT

The Forestry Research Center (FRC) has an objective of providing users with scientific knowledge by disseminating research output, so that the production and productivity of tree species improved and conservation activities are better managed.

At the moment, FRC is experiencing difficulties in fulfilling the above-mentioned objective. Among the factors for experiencing such difficulties are the isolated research, poor information and communications technology (ICT) use, lack of enough and accessible experts in the field of species selection, and lack of systematic research compilation techniques.

In an attempt to help FRC address such problems, particularly with respect to information dissemination, the present study, specifically, explores the potentiality of applying expert systems technology in species selection task. In particular, A Prototype Species Selection Expert System, SPEX, is developed by working closely with experts in the research center having years of experience and skills in the area of species selection.

The knowledge acquired from these experts is modeled using the hierarchical structure that represents concepts and parameters involved in species selection. Based on the model, the knowledge is represented using production rules. These rules are then implemented in the knowledge-pro expert system shell. Backward chaining is used in inferring the rules and extracting recommendations. The Certainty Factor (CF) for each species is calculated on priority basis to measure the belief of the human expert on the selected species.

SPEX has also been tested by the experts and users from Forestry Research Center (FRC). Based on an encouraging result/output obtained, they have established its acceptability and accuracy. The experts finally recommended that a way should be devised to build a complete species selection expert system that include all tree species, which were researched in the center.

CHAPTER ONE
INTRODUCTION

1.1 BACKGROUND

Expert systems are one of the Artificial Intelligence (AI) information systems that decrease the risks of decision-making and guide expertise in their specific subject area. Expert systems are also defined as computer programs capable of solving somewhat routine and difficult problems, usually restricted to a very specific problem area, at a level comparable to human expertise (Cercone and Mccala, 1984).

Expert systems have been applied to problems in many fields, including electronics, engineering, law, manufacturing, mathematics, medicine, meteorology and physics. MYCIN (for diagnosing bacterial infections), DART (as a computer breakdown support) and DENDRAL (assists in analyzing geological data) are some of the popular expert systems developed (Kandel, 1991).

The field of forestry is not an exception, the potentials of the expert system, thus, been explored by many developers. To elaborate the point, Schmoldt (1986) says that foresters bring a lot of implicit knowledge about the real world, awareness of local conditions and balancing of issues to arrive at their decision. This knowledge can be coded and used more or less the way the expert uses his knowledge, implying the application of expert systems in the domain.

Some of the applications of forestry expert systems reported in the literature are that of Rauscher, which is developed to recommend Silvicultural Prescriptions for individual tree

CHAPTER ONE
INTRODUCTION

1.1 BACKGROUND

Expert systems are one of the Artificial Intelligence (AI) information systems that decrease the risks of decision-making and guide expertise in their specific subject area. Expert systems are also defined as computer programs capable of solving somewhat routine and difficult problems, usually restricted to a very specific problem area, at a level comparable to human expertise (Cercone and Mccala, 1984).

Expert systems have been applied to problems in many fields, including electronics, engineering, law, manufacturing, mathematics, medicine, meteorology and physics. MYCIN (for diagnosing bacterial infections), DART (as a computer breakdown support) and DENDRAL (assists in analyzing geological data) are some of the popular expert systems developed (Kandel, 1991).

The field of forestry is not an exception, the potentials of the expert system, thus, been explored by many developers. To elaborate the point, Schmoldt (1986) says that foresters bring a lot of implicit knowledge about the real world, awareness of local conditions and balancing of issues to arrive at their decision. This knowledge can be coded and used more or less the way the expert uses his knowledge, implying the application of expert systems in the domain.

Some of the applications of forestry expert systems reported in the literature are that of Rauscher, which is developed to recommend Silvicultural Prescriptions for individual tree stands found in Minnesota, US, and PREDICT by Schmoldt that diagnoses a disease of a tree species called *Pinus resinosa* (Schmoldt & Martin, 1989).

1.2. STATEMENT OF THE PROBLEM AND JUSTIFICATION

Lack of sufficient knowledge about forests, woodlands and their components as well as insufficient awareness of the environmental role of forestry by rural communities lead to inappropriate management, utilization and conservation of tree species (Berhanu, Nyalun, and Mustanoja, 1988).

This is the background against which Forestry Research Center (FRC), previously called Forestry Research Institute, was established in 1961. FRC is currently organized under the Ethiopian Agricultural Research Organization (EARO) since June 1997 (the organizational structure of FRC is attached as Appendix I). .

Since the time it is restructured under EARO, FRC is mandated to undertake research aimed at improving the production and productivity of tree species and managing conservation activities. FRC is also mandated to provide users with scientific knowledge by disseminating research output (EARO, 1997).

The need for information in tackling the ever increasing and complicated problems relating to forestry domains as soil erosion, yield loss, increasing chemical fertilizer costs, diminishing farming lands, fuel wood and energy shortage, increasing demand for certain wood types and the like has been well established by users in this area (Berhanu, Nyalun, and Mustanoja, 1988).

Recently, FRC is working for establishment of forestry information systems with strong backing from EARO in an effort to support its research operations and the resulting provision of research output for users. However, external consultants carrying out the feasibility study for the forestry information systems project have observed the following limitations and difficulties concerning the present status of IT in FRC that might impede the continuation of the project (MOA, 1999):

- Lack of documentation.
- Non- existent integrated information system, both for management and scientific research data/ information.
- Non-operational computerized networks to facilitate resource sharing and data transfer.
- Fragmentations of scientific data that exist now on stand alone PCs with different formats and software.
- Unclear data processing policies and procedures.

Further investigation made by the researcher for the purpose of the current work, reveals the following additional problems in terms of information dissemination problems in FRC, which in fact led the researcher to the choice of expert systems in a specific task of species selection.

The main problem in the research center is the accessibility of experts in different domains. Due to the nature of forestry research, most of the researchers spend their working time away from their offices, working on field in different stations/sites around Ethiopia. This makes them inaccessible when users need them.

In addition, existing experts specialize on few uses of species (e.g., species for timber production or soil conservation etc), meaning no expert is familiar with every tree species and their uses, because it all depends which section of FRC he worked, for years. This is because every researcher works for a longer time in a specific site and with a small number of species. So, the experts with lots of experience in this case, have worked in different areas with various species accumulating a great deal of knowledge. This leads in a search for ways of transferring the knowledge of the experts which otherwise could be lost as they leave the research center.

Further, there is lack of communication among the various research sections in FRC (E.g. Silvicultural sections, social forestry section and wood utilization and research center), every section works only on its specialization creating "unnecessary decentralization of research activities" (EARO, 1997).

This led the research outputs to be isolated and individually kept in the bureaus or in the forestry journals. As a result, many sources of knowledge that a forester or forest extension worker needs currently are widely dispersed, amongst different experts or in various publications. An expert system would provide a single, integrating route into this diverse and dispersed knowledge.

It has also been observed that the researchers are disadvantaged by the incomplete and uncertain output data. This is because of the long-term nature of forestry researches and few number of research sites.

According to EARO (1997), the research in forestry is long term because research in a specific tree species needs more time to explore from few decades to a lifetime of the researcher. But poor financial background of the organization, the researcher's unsure commitment to work in the center for long years, and the pressure for immediate research output from the user and government, leads to untimely and incomplete research outputs. This makes the provision of complete information to the user difficult, and lends hands for other alternatives of information provision for users, like expert systems.

Moreover, most research components related with the suitability of species to a particular objective demand observation rather than mathematical computation. The heuristics and rules of the knowledge base can handle observational data and conclusions that are uncertain because human judgment might not be absolutely certain.

Schmoldt (1986) made it clear that only few data pertaining to a specific problem might be known for certain, especially in the field of forestry. More often the output is believed true with some subjective level of confidence.

For these reasons, there is a pressing need to design an expert system that enables forest researchers to select the right and suitable species to the specific site\place. Further motivations to the consideration of expert systems to tackle the above stated problems at FRC are summarized in the following paragraphs.

Expert systems overcome many of the limitations of the traditional procedures in forestry, because expert systems have the potential to help organize and synthesize knowledge and information of different types. They are easier to construct and also deal with uncertain information (Parsaye and Chignell, 1988).

According to Inneson (1985), since forestry inherently involves uncertainty and use of heuristics, the domain's problems are better solved by using expert systems. Numerical methods have failed because understanding about forest systems are mostly qualitative, based on experience and cannot be mathematically represented.

Expert systems are also appropriate in FRC in guiding users in the absence of experts for the specific task. The FRC senior researchers are mostly on- field working in different sites; thus, they are not available when users and junior researchers need them. But the use of expert system makes knowledge of experts working in various species selection tasks available to all users at a particular time needed.

The main beneficiaries of this research are, therefore, those researchers in forestry research center (FRC) engaging in species trials/experiments for different sites. The system is of great importance to forestry researchers because it assists them as a 'checklist'. In 'checklist' (advisory) expert system, the researcher decides as to which species have to be given priority for setting up another research for a particular area. Preliminary solutions can also be matched with expert systems; the difference in idea with the expert systems leads the user for further research and thinking about the species.

The other beneficiaries are the forestry extension workers who could use the outcome of this study as an electronic guideline for the specific domain of species selection, i.e. the extension workers use the system to check what sort of plants are recommended in their respective areas by FRC, along with further descriptions (seed treatment and storage, species names in local languages and species descriptions) needed by users. The farmers, who are in strong need of the useful and successful species in their area, are of course, the ultimate users of the system with the help of the junior researchers and forestry extension workers.

1.3 OBJECTIVES OF THE STUDY

1.3.1. General Objective

The main objective of the study is to design a prototype expert system with a view to explore the potentialities, and demonstrate the applications of expert systems technologies in a specific area of species selection.

1.3.2. Specific objectives

The specific objectives of the study include the following:

- To investigate the extent of applications of expert systems to species matching researches.
- To review literature on the different methods and techniques of knowledge acquisition and representations.
- To examine/investigate the procedures, variables and the domain knowledge/skill used in species selection process.
- To analyze and organize the domain knowledge acquired.
- To develop a prototype knowledge-based expert system.
- To test the new system for correctness and acceptability of rules as applied in Forestry Research Center (FRC).
- To give recommendations on further studies to be conducted to realize the outcomes of this research.

1.4. METHODS

1.4.1. Method of Data Collection

1.4.1.1. Interview

For acquisition and extraction of the domain knowledge and skill, interviews were conducted with the following three experts, namely:

1) Ato Mebratu Mihretu, with 23 years of experience, and head of the Silvicultural Section.
2) Ato Dechasa Jirru, with over 30 years of experience, and working in the Social Forestry Section.
3) Ato Mesekere Reta, 19 years of experience in Wood Properties Section.

Structured and unstructured interviews, in fact, series of day-to-day interviews, have been held with the above experts over a period of 6 months with the aim of acquiring the knowledge used in species selection for constructing a prototype expert system.

Unstructured interviews: - are free format interviews where the interviewer prepares only opening questions, but asks subsequent questions based on expert's answers (Zahedi, 1993). Unstructured interview methods have been used in the first sessions of interview to get a high level or general idea of the concepts confined to the particular section with respect to the species selection domain.

Structured interviews: - are types of interviews where the interviewer decides the objective of the interview, and prepares questions beforehand, and focuses the discussion on the topic (Zahedi, 1993).

The unstructured interview was a base in devising the main factors affecting species selection, through probing around the main factors. The expert is then made to draw facts and relationships from the concepts that he uses for recommendation/ decision- making.

Questions presented for experts during structured interviews are attached in Appendix III.

1.4.1.2. Literature review

For documentation of concepts and relevant techniques used in expert systems in relation to the domain area, literature, such as books, journals and the Internet have been consulted.

In addition, documents (manuals, research results, forms) were examined to assess the specific procedures and guidelines used in species selection.

The main documents used for the development of expert systems are the following: -

- **Agroforestry in Dry Land Africa, by Rocheleau, Weber and Field-Juma (1988)** - The book is used as guide by foresters for implementing forestry practices in farming communities.

- **Handbook for Tropical Forestry, by Pancel (1992)** - This is a general field guidebook for foresters in tropical regions.

1.4.2 Method of System Development

A prototyping approach has been followed to develop the system using the technique of Certainty Factor (see section 2.5.2.3. for detailed description). Prototyping is preferred from the other methods, because it is a common way of developing the expert system (Parsaye and Chignell, 1988). The other reason for adopting prototyping is that it enables better management of the system in times of implementation – i.e. testing and demonstration.

The Certainty Factor (CF) is preferred for developing the prototype due to (Masuch, 1990): -

- Simplicity and ease in handling complex or combination of rules or facts.
- In most cases, the CF yields acceptable results.
- Large number of expert systems prefer the CF. Large number of expert systems, as MYCINE (expert medical diagnosis system) and PREDICT (species diagnostic consultation expert system) prefer the CF.

Lastly, the Prototype Species Selection Expert System, SPEX, creates an easy-to- use question and answer session for the user to decide and give recommendations on the best species to plant on the area, thus demonstrating the expert system's applications in providing easily accessible, integrated and uncertain information.

Programming tools - The knowledge-Pro version 3.0, the expert systems shell that is available in SISA, has been used for developing the expert system. Knowledge-Pro has debugging tools; design tools for point and click design, and flexible help system. Moreover, the code is readable and easier to update and maintain. It is also a multipurpose shell capable of deciding in a wide variety of knowledge domains (KpWin, 1991).

Testing - was done mainly by experts and users by going through the content of rules and acceptability of the recommendations, as part of validation and verification of the prototype. Then to show whether the system proposes the right species on priority basis, using a Certainty Factor, and to make the user comment on the result, a test data prepared by experts is fed in to the system.

1.5 SCOPE AND LIMITATIONS

The scope of the study is exploring the potential of using the expert systems in specific task of species selection. Because of time limitations, the research is focused on modeling and developing a prototype expert system. So, further elaboration and criticism of experts and users, as well as iterative development process is needed to come up with a full-fledged expert system, which is beyond the scope of the present study.

Other artificial intelligence systems, like artificial neural networks (ANN), Bayesian belief networks (BBN) and genetic algorithms (GAs) were not considered in the choice due to lack of resources, like costy and complex Soft wares / System shells and need for more data. This made the study to be limited to the Certainty factor (CF) technique and the expert systems in general.

Also some other AI systems were not considered because these systems could not fit into the subject matter. For instance, fuzzy expert systems are categorized under those techniques handling vagueness rather than uncertainty.

Additionally, due to time limitation, representative species have been chosen as a sample for testing purpose. Otherwise the number of woody plant species (trees and shrubs) found in Ethiopia exceeds 300 species, treatment of all these species needs a lot of time, hence, it is beyond the scope of the current work.

1.6 ORGANIZATION OF THE THESIS

The present study is organized into six chapters. Chapter one is the introductory part, which contains background, problems and justifications to conduct the research, objectives and methodology to carry out the research, and its scope and limitation. Review of literature on the expert systems, about its background, architecture and expert systems applications in the area of forestry is presented in chapter two.

Chapter three focuses on the plantation information in Ethiopia. The chapter includes the background history and types of plantation activities, along with the reasons behind administrating plantations for the sites and how the species selection process and research assist in having a successful plantation activity.

Chapter four discusses the methodologies used in knowledge acquisition sessions including the interview, and knowledge structuring and modeling processes.

In chapter five, the development and integration of the prototype expert systems has been discussed using a sample demonstration of the user interface. The last chapter, chapter six, presents recommendations accompanied by concluding remarks.

CHAPTER TWO
EXPERT SYSTEMS

2.1 INTRODUCTION

Since the late 70s, expert systems emerged as the most successful application-oriented branches of Artificial intelligence (AI). According to Brule (1986), AI is a sub- field of computer science, aimed at producing computer systems that exhibit 'intelligent' action in solving variety of problems such as planning, selection, interpretation and diagnosis.

Thousands of expert systems are now in routine use worldwide and span the full spectrum of activities in business, agriculture, medicine, industry and government. This is because the primary goal of expert systems is to make expertise available to decision makers, researchers and technicians who need answers quickly. Moreover, there is never enough expertise to go around; certainly, it is not always available at the right place and at the right time. Secondly, expert systems mostly rely on logic, belief, rules of thumb, opinion and experience, which was becoming the indispensable source for those disciplines mathematics could not defeat (Feltovich, Ford and Hoffman, 1997).

A good readable account on expert systems, i.e., their definition, components, and distinguishing features from other systems, are provided in this chapter. Related concepts on types and techniques of handling uncertainty, Certainty Factor, the approach used in this thesis work are also discussed in detail. Towards the end of the chapter, some applications of expert systems in forestry domain are reviewed.

CHAPTER TWO
EXPERT SYSTEMS

2.1 INTRODUCTION

Since the late 70s, expert systems emerged as the most successful application-oriented branches of Artificial intelligence (AI). According to Brule (1986), AI is a sub- field of computer science, aimed at producing computer systems that exhibit 'intelligent' action in solving variety of problems such as planning, selection, interpretation and diagnosis.

Thousands of expert systems are now in routine use worldwide and span the full spectrum of activities in business, agriculture, medicine, industry and government. This is because the primary goal of expert systems is to make expertise available to decision makers, researchers and technicians who need answers quickly. Moreover, there is never enough expertise to go around; certainly, it is not always available at the right place and at the right time. Secondly, expert systems mostly rely on logic, belief, rules of thumb, opinion and experience, which was becoming the indispensable source for those disciplines mathematics could not defeat (Feltovich, Ford and Hoffman, 1997).

A good readable account on expert systems, i.e., their definition, components, and distinguishing features from other systems, are provided in this chapter. Related concepts on types and techniques of handling uncertainty, Certainty Factor, the approach used in this thesis work are also discussed in detail. Towards the end of the chapter, some applications of expert systems in forestry domain are reviewed.

2.2 DEFINING EXPERT SYSTEMS

The definitions of expert systems vary from author to author. This is because definitions are mostly based on functions (*what* the system does), while the others are based on structures (*how* the system does). These definitions are given below assuming the rule-based rules.

- Expert systems solve problems that are difficult enough and require significant human expertise for their solutions (Feigenbaum and Corduck, 1983).
- Expert systems are programs that mimic the advice giving capabilities of human experts (Brule, 1986).
- Expert systems are representations of domain specific knowledge in a manner which the expert thinks through the incorporation of ... ways of handling uncertainty (Liebowitz, 1988).
- Expert systems are intelligent computer programs that use knowledge and inference procedure (Harmon and king, 1985).

The first two of the above definitions employ functional components of expert system, while the other two are based on structural definitions. However, Inneson (1985) gives a comprehensive definition that combines both functional and structural components as shown below:

> *An expert system is the embodiment within a computer of a knowledge-based component from an expert skill in such a form that a system can offer intelligent advice or take an intelligent decision about a processing function.*

2.3 COMPONENTS OF EXPERT SYSTEMS

Expert systems have four main components, the knowledge base, the inference engine, explanation facility and the user interface. These components have been summarized and modified from Stefik (1998) in Figure 2.1 below, followed by their brief description.

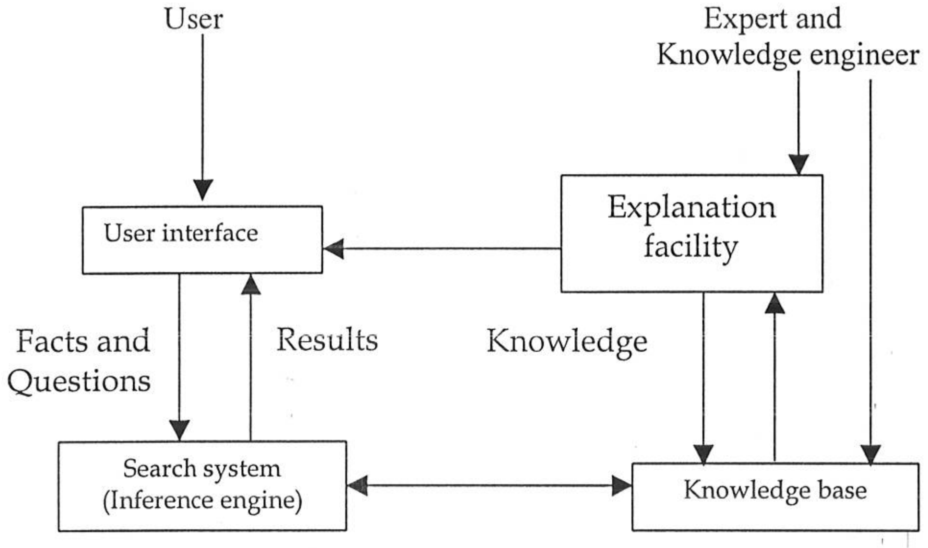


FIGURE 2.1. THE CLASSICAL CHARACTERIZATION OF THE KNOWLEDGE BASED SYSTEMS.

2.3.1 Knowledge Base

Knowledge base holds the procedural and implicit knowledge specific to the problem area that the system is set up to solve; it is used by the system to reason about the current problem (Alison, 1994).

To represent knowledge into a knowledge base, different knowledge representation methods are available. Among the many, semantic networks, frames and rule-based systems can be mentioned as popular methods (Zahedi, 1993; Ebrahim, 1999). The three methods are briefly described in the following paragraphs: -

- **Semantic Networks** - consist of nodes and arcs. Each node represents concepts and each arc represents a relationship between the two nodes it connects. The networks are suitable to modeling complex structures such as language structures.

- **Frames** - are collections of knowledge about an entity or concept. Each frame is a complex package of slots, which are data values describing the many characteristics or attributes of an entity, including its relationships with other entities. However, constructing knowledge using frames is more complicated and costly.

- **Rule-Based System** - One of the widely used strategies/approaches used for Knowledge base representations of the domain knowledge is the production rule (or simply rules).

A rule normally consists of IF' (condition) and 'THEN' (action) parts. Rules involve proven facts (facts widely shared within the domain) and hypothetical facts (facts inferred based on individual's experience). The latter provide a means of using coefficients of certainty (Pederson, 1989).

For instance, IF light is coming on
AND NOT engine is turning on
THEN problem is with starter.

So, the rule states that if the two conditions/hypothesis are true, then the system recommends that the problem be on starter. A rule-based system is used in this study to represent the domain knowledge base,

2.3.2 Inference Engine

The search system (an inference engine) is another component of expert system, which executes the knowledge base by carrying out set of actions that utilize the knowledge in finding a solution to the problem.

The inference engine uses symbolic information and knowledge in the knowledge base to form a line of reasoning in solving the problem (Drenth, Morris and Gweyneth, 1991). The two types of reasoning used to control and organize the steps taken to solve problems are forward and backward chaining.

- **Forward chaining** - which is also called data-driven chaining, is the general term used to describe a group of search strategies that starts with set of conditions and try to infer the conclusions implied by those facts. For example, in the following rule,

  If rainfall is <=300 (premise of the rule)

  Then zone is dry land (conclusion of the rule).

According to Parsaye and Chignell (1988) the forward chaining is used especially when the objective is to flash out all facts. The domain types that forward chaining works range from design and planning, to interpretation.

- **Backward Chaining** - which is also called goal-driven chaining, is used to describe a group of search mechanism where only information that helps solve the problem is asked, which leads to short and focused consultations (Feigenbaum and Corduck, 1983).

2.3.2 Inference Engine

The search system (an inference engine) is another component of expert system, which executes the knowledge base by carrying out set of actions that utilize the knowledge in finding a solution to the problem.

The inference engine uses symbolic information and knowledge in the knowledge base to form a line of reasoning in solving the problem (Drenth, Morris and Gweyneth, 1991). The two types of reasoning used to control and organize the steps taken to solve problems are forward and backward chaining.

- **Forward chaining** - which is also called data-driven chaining, is the general term used to describe a group of search strategies that starts with set of conditions and try to infer the conclusions implied by those facts. For example, in the following rule,

  If rainfall is <=300 (premise of the rule)

  Then zone is dry land (conclusion of the rule).

According to Parsaye and Chignell (1988) the forward chaining is used especially when the objective is to flash out all facts. The domain types that forward chaining works range from design and planning, to interpretation.

- **Backward Chaining** - which is also called goal-driven chaining, is used to describe a group of search mechanism where only information that helps solve the problem is asked, which leads to short and focused consultations (Feigenbaum and Corduck, 1983).

A backward chaining usually works well for problems in which the explanation for candidate solutions is pre enumerated, i.e., the solution to some problems are known beforehand. Here is an example (Stefik, 1997): -

An auto repairman knows that cars could not start for one of several possible reasons. The answer is chosen from a relatively small group of common causes that the expert is familiar. So, his job is to track which one of the problems is causing the car failure. The expert then reasons 'backwards', asking questions whose answers will help him match those causes to the symptoms for this particular car.

Backward solution usually works for selection, classification, and diagnostic problems (Parsaye and Chignell, 1988).

The following table compares and contrasts the two inference chaining methods (Parsaye and Chignell, 1988; Zahedi, 1993):

TRAIT	FORWARD CHAINING	BACKWARD CHAINING
Solutions	Not preenumerated	Preenumerated
The goal is:	Not necessarily known	Known
The objective is to	Flash out all facts	Infer one key fact
Starting situation	Some facts known	Few facts known
Q&A type input required from user	No	Yes
Strategy	Build solution	Detect solution

Table 2.1. Questions relevant in choosing between two search strategies

2.3.3 Explanation Facilities

Explanation Facilities Provide detailed descriptions of the reasoning behavior of the expert systems during consultation, allowing the user to inspect selected parts of the inference.

This is because expert system uses heuristic knowledge where, like humans, its credibility is often questionable. Moreover, people tend to know the rationale and the extent of plausibility of expert systems' consultations. The explanation facilities provide the explanations for the reasoning to ascertain plausibility of the recommendations. Explanations can be generated by tracing the line of reasoning used in the inference engine (Benchimol, Levine and Pomerol, 1987).

Parsaye and Chignell (1988) explain about the two main types of explanation facilities in detail. Here are some of the discussions: -

- **'WHY' Facilities** - allow the user to examine why a certain question is being asked. It shows which rules the inference engine is currently using. Concerning the WHY facilities the following example is given by Stefik (1998):

  The identity of the organism is e-coli (0.74).

  Why?

  The following rule was used for the conclusion.

  Rule 156: IF site of culture is blood

  •
  •
  •

  THEN there is suggestive evidence (0.74) that the identity of the organism is e-coli.

- **'HOW' Facilities** - offer means of investigating how values for a particular attribute should be established, i.e. which rules have been applied for deriving its values. An example from MYCIN'S explanation facility goes like this: -

(Recommendation by the system) culture-1 was from a sterile source.

(Input from the user) HOW?

Rule 189 is used to conclude that this blood culture was taken from sterile source. This gave a cumulative CF of 1.0.

2.3.4 The User Interface

User interface gives the user a chance to interact with the system. It is the critical part of the expert system to be designed with minimum of errors, because the user is basing his judgments in the results obtained from the system. By utilizing the user interface, questions are presented for the user, where the user answers the choice given by the system or on the space provided, and gets an advise/ recommendation from the system (Richardson, 1995).

2.4 FEATURES OF THE EXPERT SYSTEMS

There are some distinguishing features of expert systems from the human expert, and from conventional programming languages: -

2.4.1 Expert Systems Vs. The Human Expert

A human expert is the one that can draw upon a comprehensive knowledge about a very specialized field or domain. An expert is also capable of solving a particular problem, structuring information and knowledge in such a way that it usually divides a problem into smaller, more easily solvable parts. An expert can also find multiple solutions, if they are appropriate, and can justify, verify or at least attach some level of certainty to its solution (Alison, 1994).

The expert systems incorporate attributes discussed above for the human expert. In addition, expert systems display the following characteristics over a human expert (Englemore, 1993):

- **Consistency** - expert systems make comparable recommendation for like situation. But according to Brule (1994) humans are influenced by:
  - Recency effects (most recent information have disproportionate impact on judgment).
  - Primacy effects (early information dominates the judgment).
- **Reproducibility** - expert systems can be duplicated and made in many copies, whereas it is expensive and time consuming to train human experts.
- **Breadth** - the knowledge of multiple human experts can be combined to give the system more breadth than a single person is likely to achieve.
- **Transparency** - expert systems also enable users to see why decisions have been reached.
- **Knowledge Acquiring** - expert systems can acquire knowledge from several experts within and between subject areas more than the humans do.

2.4.2 Expert Systems Vs. Conventional Programming Languages

What distinguishes an expert system from ordinary computer programs, according to Yaghamai et al (1984) is that in conventional programs, like Basic, Pascal and FORTRAN etc., pertinent knowledge and methods for utilizing them are all intermixed. In an expert system, the problem-solving model appears explicitly as a knowledge base rather than implicitly as part of the coding, and the knowledge base is manipulated by a separate and clearly identifiable control strategy, and an inference engine.

Therefore, unlike conventional computer applications, expert system incorporates data into the instructions in such a way that new knowledge can be added to the program without extensive reprogramming (Zahedi, 1993).

In addition, conventional programs tell nothing about the overall functionality of the program. The program is simply the instruction script telling the computer what to do. While expert systems set free of the intermixing, and divide the task into the database of, i.e., knowledge base, and inference engine which uses several methods of 'inferencing' to select which rule in the knowledge base to process.

A feature making an expert system different from other system developments is the preconditions that must be fulfilled. Failure on one of the series of preconditions to develop expert systems might jeopardize the development process. Some of these criteria are listed below (Combs, 1984; Pederson,1989):-

- Getting genuine expertise using judgment and experience.
- Willingness of the experts for the study.
- A need for focus in a narrow area/ specialization.
- The inadequacy of research experts in the area.
- A need for consistent and undisputable subject area.

2.5 UNCERTAINTY

In most fields of science, like agriculture and medicine, knowledge is for the most part inexact and uncertain. People make decision in an environment where facts and rules make various shades of vagueness, imprecision, and errors. A number of methods have been proposed to superimpose uncertainty over the logical structure of expert systems (IPT, 2000).

2.5.1 Kinds of Uncertainty

There are various types of uncertainty, as described by Pederson (1989):

A) Probability

One common way of thinking about uncertainty is in terms of probability. Probability is usually expressed as a chance of some particular state occurring as a percentage of population of all possible states. Probability is useful when we have a model of how the system works. For example, we know the coin is weighted proportionally, has equal area on both sides, and thus either side has equal chance of appearing.

B) Statistical uncertainty

The other way of determining uncertainty is calculating the behavior of, and the certainty of various results statistically by examining past behavior. Such uncertainty determinations rely on empirical evidence (personal judgment).

C) Heuristic uncertainty

Heuristic uncertainty is uncertainty derived from heuristic knowledge, which is less vigorous, more experimental and more judgmental knowledge of performance. Such knowledge is expressed by using rules of thumb or words like very likely, high, low etc. Heuristic knowledge underlies uncertainty, because it is expressed as the "art of good guessing" (Englemore, 1993).

2.5.2 Handling Uncertainty

Representing uncertainty is one of the most controversial subjects in expert systems. Expert systems differ from conventional software because of their human like reasoning. Therefore, it seems sensible to expect the expert systems should be capable of handling uncertainty (Zahedi, 1993).

A similar character that all the representations of uncertainty share is that regardless of the techniques available in dealing with uncertainty (i.e., determination of likelihood, probability, degree of truth, certainty or confidence factor...), the techniques may be essentially viewed as the process of selecting a point on a scale, say, a number between 1 and 100 (AAAI, 2000). As the number approaches 100, the expert's belief on a given statement is increasing.

In handling uncertainty, there are a number of methods used. These methods and their differences are briefly discussed below.

2.5.2.1 PROBABILITY APPROACHES

I) Subjective (Bayesian) probability

The most accepted probability approach for expert systems is the subjective (Bayesian) approach.

In the Bayesian probabilities, p (h_i)(the probability of the hypothesis i) is called the prior probability, and p (h_i/e)(the probability that hypothesis h is true given the evidence e) is called posterior probability. The prior is the expert's belief in the truth of the hypothesis, whereas, the posterior is the revised belief of the expert, after observing an event (Stefik, 1998).

Normally, the application of the Bayesian approach is better understood by depicting the rules of the inference process in graphical form as belief or inference network. A belief network is a graphical network that represents probabilistic relationships among variables (AAAI, 2000).

For example, in a simple belief network, where the three variables are M, C, R, as shown in the figure, each variable has two elements. M represents the question of whether the meat is missing, C represents the cat ate the meat from the kitchen, and W represents the question of whether the cook used the meat in making 'wat'.

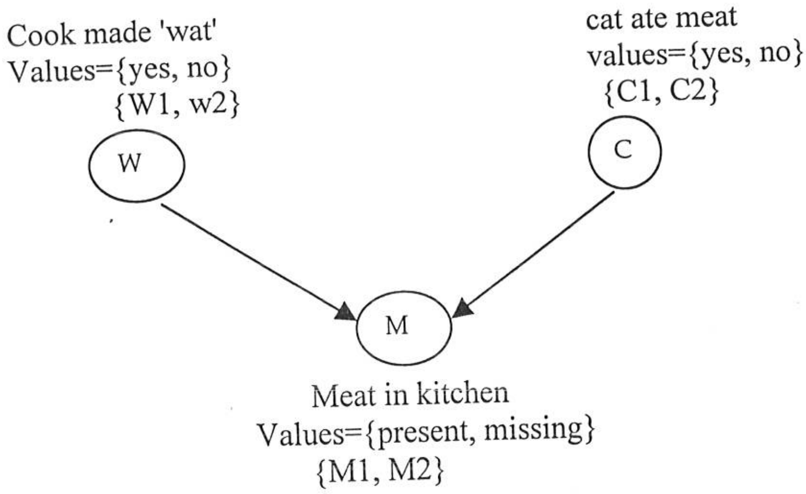


FIGURE 2.2 THE BELIEF NETWORK FOR THREE VALUES, M, W, AND C.

Then, the belief of the observer is put as a probability coefficient for each element of the variable. For example, M2={0.4} is the observer's belief that the meat in the kitchen is missing is 0.4.

Some of the limitations of the Bayesian theory are (Zahedi, 1993): -

- For real world expert systems with many rules the Bayesian approach needs too many probabilities. Shortliff, as quoted by Parsaye and Chignell (1988), pointed out that vast proportions of areas such as medicine suffer from having too few data and so much imperfect knowledge that a rigorous probabilistic analysis is not possible.

- Baye's law is mathematically correct only if all possible outcomes are disjoint. Although one may reformulate concepts, the results often don't confirm/respond to the intuitive concepts of the expert.

i) Objective probability

In the objective or frequentist interpretation of probability, p (h) -the probability of belief in the hypothesis h represents the actual or theoretical frequency of a hypothesis being true. In this approach, probability values should be measured by observing the events, or by theoretically deriving them from the underlying model of the system (Parsaye and Chignell, 1988). For example, if a coin's toss of 6 times resulted in three heads and three tails, then, the interpretation is that a coin toss generally results a head 0.5 times.

2.5.2.2. Dampster- Shafer (d-s) theory of evidence

The D-S approach is a calculus for numerical degrees of belief. It was first described by Dampster and later extended by Shafer in 1960. The approach attracts expert systems communities because it side steps requirements for completeness in probabilistic information, and because it is compatible with formalism of logic and dbase query languages.

According to Stefik (1998) the d-s theory of evidence, the ground set of elements are denoted by θ, and θ is considered to be mutually exclusive (i.e. no two hypothesis can be true at the same time) and exhaustive. For instance, for θ={a, b, c}, then if 'a' is measles and b is chickenpox, then {a, b, c} is " measles, chickenpox". All the other measures are based on M, the basic probability assignment and sum of assignments over all subsets is 1.

Some of the limitations of d-s theory are: -

- The d-s theory is not strictly based on probability theory and could produce results that conflict with people's intuitive notion about how to measure belief.

- The theory also assumes the n- hypothesis that forms the environment are mutually exclusive, which are not the same in the real situation (Zahedi, 1993).

2.5.2.3. Certainty factors (CF)

Due to the limitations of the above systems of handling uncertainty, some expert systems have used certainty factors. Certainty factors, in the handling of uncertainty, are to a great extent empirical.

Certainty factors (CF) are attribute or rule properties used to describe the level of certainty or confidence in a particular value. Pederson (1989) asserts that CFs is used to represent the certitude of user input, assertions, and rule conclusions. They are numeric values maintained by the inference engine, usually stored as values ranging from 0 to 10.

The CF model has been developed to introduce the concept of uncertainty in the medical diagnostic expert system (MYCIN). The model is also used in large number of expert systems. Its success stems from its simplicity, i.e., CF is easy to handle and in most cases yield acceptable results (Masuch, 1990).

Although certainty factor model is inspired by probability theory, it deviates from probability theory in certain respects. One of the deviations of CF is its approach in sidestepping the weakness of probability theories of lacking sufficient data to accurately estimate prior probabilities, by using expert estimated certainty factors, instead of measured frequencies.
The other method used in CF, but stands in sharp contrast with other probability theories is the combining and updating technique of CF. This technique employs three methods (Stefik, 1998): -

- Locality- methods of combination apply regardless of what else is present.
- Detachment- facts can be used for further inference regardless of how they were derived.
- Modularity- combination of locality and detachment in logic.

The above characters of CF are better described by the properties it displays in combining CF. There are two ways of combining rules in CF (Stefik, 1998; Pederson, 1988): -

A) **When the two rules have the same conclusion**, then, the rules are combined by adding using the following formula. Both A and B are hypotheses of rules:

Combined CF =(A+B)-(A*B)

Where,

A= the confidence in Hypothesis (H) based on the evidence, which is Rule 1

B= CF (H, rule 2) i.e., the confidence in Hypothesis (H) based on the evidence, which is Rule 2.

B) **When conclusion of one rule is hypothesis to another rule,** then, the following alternative formula is used: -

Combined CF = A*B

Where,

A= the conclusion CF for rule 1, which is now the hypothesis of the current rule 2.

B = the conclusion of rule 2, which is updated as its combined to A.

The later technique of updating CF is used in this study, due to the nature of the rules following one another, i.e. as explained briefly in chapter 5, the conclusion of the rule to get the site, is used as hypothesis for the rule to get the recommended species. Such updating/combining of CF is carried out to ascertain the strength of evidence, in times when the conclusion of one rule becomes the evidence for the next rule in the chain. In these chains, the certainty associated with a piece of evidence must enter into the certainty equations for further conclusions.

Mathematically speaking, let H be a hypothesis supported by evidence, E, and H is the hypothesis that can be derived from E` by a given rule. If CF` is the weight that would be used in the case that E` were known with certainty, then CF gives the revised weight. Restated, the CF of the rule is reduced simply **by multiplying** it by the known weight of the evidence (Stefik, 1998). The formula is given as:

CF(C) = Y * (MAX (0, CF (B))). = Y * X

For example, IF A, THEN B (0.2)

IF B, THEN C (0.5)

The revised CF for C is 0.5* 0.2 =0.1, because the hypothesis B has lower CF, then, the resulting CF or belief of the expert as a result, decreases.

2.6 EXPERT SYSTEMS APPLICATIONS IN FORESTRY

2.6.1 Expert Systems in Orange Fruit Tree

The expert system for managing orange fruit tree (CITEX) has been developed in the Egyptian Central Laboratory for Agricultural Expert Systems, CLAES (1996) in 1997. CITEX consists of three subsystems namely site assessment, agriculture practice management, and disorder diagnosis and treatment.

The main function of the site assessment subsystems is to generate one of these decisions: the site is perfect for cultivation, a set of treatment operations has to be applied to enhance the soil and water characteristics, or the site is not suitable for cultivation. And disorder diagnosis and treatment subsystems are the same as the corresponding subsystem of CUPTEX (expert system for managing cucumber production). CITEX was developed using NEXPERT/OBJECT commercial shell and then ported to KROL (tool-based on logic programming and object oriented programming paradigms) as CUPTEX.

The methodology is divided into two main parts: knowledge engineering, and software engineering. The two parts of the methodology are interacting through a spiral model. The output of the knowledge engineering process is fed as input to the software engineering activities. The main interaction is through feeding the results of the knowledge analysis phase into the design activity. The methodology also preserves the design model through introducing a method to transfer the design into an expert system development shell.

At last, to validate the experiment in each subsystem, a performance factor is calculated for each of the best three human experts chosen and the CITEX expert system, using the score system. The result showed that except for the irrigation subsystem, CITEX is the best performer and it has the best overall system performance. That is owing to the score of 3.16 out of 4, to CITEX, against the best performing expert, 2.13 and 2.79.

2.6.2. Expert Systems for Denitrification

Inneson (1985) produced an expert system, which has the objective of predicting under what environmental conditions (levels of soil moisture, soil nitrate, temperature, rainfall rates) the process of denitrification (when nitrogen has been lost as gas) occurs to any significant extent at a specific soil in Berkshire, UK.

By taking a data where statistical approach failed to explain the relation between rates of denitrification and environmental conditions, Inneson (1985) applies a rule-based approach to translate the data into machine-readable form.

The rule inference is done by HULK (user written software), which is suitable to deriving rules from the data and to screen the independent variables (soil temperature, soil water, precipitation, soil nitrate) to identify those that appear linked to the dependent variable (denitrification) under investigation.

The testing of the expert system was made by LEAP (user written software), which measures the effectiveness of the rule in satisfying the hypothesis, in this case Denitrification is greater than 20. The result showed that a 100% success was achieved. The expert system then, recommended that complex environmental systems could be built in such a form to facilitate our understanding of how the natural systems work.

2.6.3. Predict: The Pest Diagnosis System

PREDICT *(Pinus Resinosa Expert* Diagnostic Consultation Tool), was developed by Schmoldt and Martin (1989). The authors decided to implement the pest diagnosis system under two different shells, EXSYS and INSIGHT2+, to see if the performance of the final system would be influenced in any way by the particular shell used.

Both EXSYS and INSIGHT2+ use a rule-based method of knowledge representation, i.e., knowledge is represented in the form of IF-THEN rules, and decisions are made by drawing inferences from these rules and the characteristics of a specific problem in pest diagnosis of *Pinus resinosa*.

Both development tools also employ a backward-chaining control strategy. A logic and completeness rule block was also constructed to deduce facts omitted by the user and to minimize the need for questioning.

The primary objective of this study were to develop an expert system for pest diagnosis of red pine that would be useful for field foresters with little formal training in forest pathology or entomology. The secondary objective was to determine if the expert system could perform on a par with recognized experts, i.e., forest pathologists and entomologists.

As a result, PREDICT recognizes 28 damaging agents including species of mammals, insects, and pathogens, as well as two types of abiotic damage.

Input to PREDICT is obtained from pest damage reports containing specific information about stand/site conditions, tree symptoms, and signs. Diagnoses from PREDICT take the form of a list of one or more possible agents with corresponding confidence values. It was necessary to develop special procedures for refining and evaluating the system to accommodate the often vague and uncertain nature of pest damage information.

Two versions of PREDICT with two shells were evaluated and compared with three recognized experts and two field foresters. No significant differences were found between the performances of PREDICT and the experts; however, PREDICT performed significantly better than the two foresters, even though they both have training and experience in forest pest diagnosis (predict versions scored 7.45 and 6.8 out of 10, but the two foresters scored 5.05 and 5.6 out of 10 in performance). It was concluded that PREDICT is able to improve the diagnoses of field foresters to a level comparable with recognized experts.

CHAPTER 3

PLANTATIONS INFORMATION IN ETHIOPIA

3.1 INTRODUCTION

Ethiopia is located in the horn of Africa between latitude 3° and 14° N of the equator and latitude 33° and 48 °E. Its diverse physical and environmental factors have resulted in a great diversity of tropical, subtropical and temperate climates.

The large diversity of ecological conditions is determined by topography, ranging from 110m below sea level at Afar depression to the pick of Mt. Ras Dejen, 4620m. The environmental diversity brings about variation in flora and fauna. The number of woody plants is around one thousand, out of which 300 are tree species (MOA, 2001).

Forests and trees are essential for the country's economic and environmental security. These resources contribute significantly to Ethiopia's economic development, preservation of bio-resources and the protection of environment. So, the challenge for the researchers in forestry is to find a way to arrest the loss and ensure sustainable utilization through engaging in plantation activities with the appropriate and accurate information at hand.

In this chapter, history, types and reasons behind plantation activities of tree species in the country are briefly described. The main prerequisite to have a private or community plantation projects, i.e. species selection process and the research going on in the Forestry Research Center (FRC) are also briefly described.

3.2 HISTORY OF FOREST PLANTATION IN ETHIOPIA

Looking back to the world plantation situation, though native trees were used for a wide range of plantation purposes, such as shelters, fuel wood, weapons etc., it was not until the 19th century that scientifically managed plantations of native and exotic species were established (Boland, 1985).

Extensive exotic plantations of eucalypts and pines were planted widely during the 20th century for the industrial wood, and this development overshadowed the need, especially for developing nations to cultivate trees for other purposes as fuel wood. In the same token, first large-scale plantations were held in Ethiopia through the introduced Australian tree species in 1895 to solve the problem of severe shortage of fuel wood and poles around the capital city, Addis Ababa (Boland, 1985).

Berhanu, Nyalun and Mustanoja et al. (1988) explains that Mondonvidallet, a French philologist, is to be credited for the first introduction of eucalypt and acacia species and their eventual large-scale production. Of these species, E.globulus has done remarkably well and has been extensively planted in towns and villages of the western plateau. At lower elevations, E.camaldulensis plays a similar role.

The large-scale plantations were further strengthened with the trial/experimental plots and an arboretum, established in early 60s. The earliest Forestry Research Section of the Chilalo Agricultural Development Unit aided by the Swedish International Development Agency began experiments/trials in 1967 introducing 24 eucalypts. Most of the introduced species are obviously well adapted to the climatic and edaphic features of the country (Husnia, 1997).

There was realization by 1970s of the unfavorable impact that the exploding world population was having on woody vegetation. Excessive cutting of fuel wood, overgrazing and mode widespread cultivation have led to series degradation of the native tree flora in many areas.

Due to the above reasons, tree planting considering its adaptability in an area and various end-used becomes more and more a necessity in developing countries in general and Ethiopia in particular.

3.3. TYPES OF PLANTATIONS

In Ethiopia, Plantations are mostly carried out and established by the government, community woodlots or catchments/protection plantation. Community woodlots/plantations (managed by the farmers or a community) can either be Protection-Oriented or Production-Oriented; whereas, Peri-urban (located around Addis Ababa and major towns) and industrial plantations are exclusively production directed.

EARO (1997) has briefly explained the two types of plantations as follows: -

- **Protection-oriented plantation** - this is the type of plantation operation used to prevent land degradation (soil and vegetation loss) and having the effect of increasing the fertility of the soil through the years. Such operations take place in sloppy and badly degraded lands as well as around catchments for protecting the soil from erosion.

- **Production-oriented plantation** - this type of plantation is mainly concerned with producing fuel wood and timber production. Some of the other purposes of Production-Oriented plantation are animal feed, farm boundary planting and intercropping (planting a tree side by side with crops) for crop yield increment.

3.4 REASONS FOR PLANTATIONS

The aims of plantations of tree species, done mostly by introducing species into a locality might be to meet a particular need for forest goods and services which can not be economically or satisfactorily derived from the indigenous (locally available) forest species. To assess whether need for plantation exists in the country, we shall look into some points in a brief manner:

3.4.1 Insufficient Local Demand for Fuel Wood and Timber Products

A. TIMBER PRODUCTION

Wood/Timber can be produced into poles, posts, lumber, pulp and paper production, construction, furniture and the like. Though precise figures are not available, it is estimated that about 24 million m^3 of wood is produced annually, of which 10% is only used for industrial and building purposes, and the reminder for fuel wood and charcoal (Berhanu, 1997).

In the country, where there are more than 250 different forest products processing and joinery industries using very limited (non-) forest products or resources, critical shortage of raw material supply can be manifested. Besides, among more than 320 different tree species that can be of use to a variety of forest products for different purposes, only a few selected indigenous species were utilized (WURC, 1995). The statistics shows that amongst the end-uses of timber, especially Poles and Posts are now in short supply by 1.3 million m^3 and sawn wood by 351,000m^3 (EFAP, 1994).

This increase in demand for wood in which the indigenous natural forests could not keep in pace, and creating more forest depletion, prompts a search for exotic species which are fast growing, and with high volume /yield per tree.

According to Husnia (1997), the demand for forest products is increasing because of increases in population and the standard of living. Consequently, larger and larger proportions of the forest areas are devoted to intensively managed and fast-growing exotic species of eucalypts, acacias and casuarinas. Between 1974/75 and 1983/84 increasing areas of introduced and indigenous species were planted for fuel wood, poles, and pulp and paper on 200,000ha area all over Ethiopia, of which 65% were peasant plantations and the rest, government plantations.

B. FUELWOOD PRODUCTION

Fuel wood and tree residues consist of a large portion of (70%) the energy source in rural Ethiopia. But the supply (20 million m^3) could not meet the demand of fuel wood energy from mostly rural population (49-64 million m^3) even by half, indicating significant shortage of supply (MOA, 1992). This prompts the users and researchers to find another way of easing the shortage, by researching on the other equally important but little known and under-utilized species.

3.4.2 Multipurpose Species Plantations

Multipurpose species are used as a natural way of coping with an ever decreasing agricultural land and the longer years a singletree species occupies the area compared to agricultural crop.

According to Cremer (1990), agro forestry (social forestry) practices (tree planting in and around farm land) help by introducing multipurpose species, which are able to harmonize the decreasing land and the increasing need for food, fuel, cash income, fodder (food for animals), and building materials.

3.4.3 Deforestation

Deforestation is the removal of vegetative cover which reinforces a process of land degradation causing agricultural yields to fall, and eventually leading to loss of fertile land (MOA, 2001). Nowadays, deforestation is increasing in size and extent. The forest cover recently reached an estimated less than 2%. The current annual loss of high forests estimated at 150,000- 200,000ha. The reasons for deforestation, for purposes of clearing for agriculture, grazing, cutting for fuel wood etc, are closely related with the viscous cycles of factors i.e., population growth, state of the environment and poverty (Berhanu, 1997).

3.4.4. Soil Reclamation and Fertility

Environmental degradation that is caused mainly by deforestation, involves both soil erosion and loss of soil fertility. According to EFAP (1994), in the early 90s, about half of the arable/cultivable land is seriously eroded. The impact of soil erosion is displayed through the crop production with a 1-2% loss per yield. At the same time, 20-30 thousand ha of cropland in the highlands is abandoned annually. This calculation would lead to conclude that by 2010, the erosion destroys the land of about 10 million farmers. Additionally, reduced infiltration of water into the surface of the soil affect the availability of water for human use all over the year. This has consequences of flood damage from the rivers /streams.

Choosing the right species in alleviating the previously discussed problems is indispensable and is the only less costly and high yielding problem solving mechanism held both by forest managers and researchers (Pancel, 1992).

3.5 SPECIES SELECTION RESEARCH: AS NECESSITY FOR PLANTATION

All planting programs require clearly stated objectives before planting commences with all economic staff and time–frame constraints being carefully analyzed and documented (Pancel, 1992). As one of the planting programmes, species selection directly determines the success of the plantation activity. Having decided upon a certain species, the project is committed to a product type(s) for at least a rotation period.

Naturally, the indigenous species already thriving in the project plantation area will meet the objectives of the plantation plus the site conditions from the safest choice. But for the matching of requirements for exotic species in a new plantation area, adequate information on species requirements of the site and climate, natural range, adaptability, the proposed end-use, and other inherent qualities of the trees determine the success or failure of the species to be introduced.

On the other hand, the selection of species is still not a precise science and largely reliant upon personal judgment and experience augmented by literature reviews (Boland, 1985).

Few individuals or groups have such required knowledge, and there is always an inexperienced forestry staff in the area. So, reliable information on ecological, silvicultural and utilization characteristics of many potentially valuable species is still unavailable and hampers species selection.

But in Ethiopia, species selection information and recommendations have been consulted late after the plantations have been taken, which Pancel (1992) has proved in his research after examining 42 reforestation projects of which 60% received additional species information during their life-span that resulted in a new species selection.

According to Pancel (1992) regarding the objectives of the species- selection, the following could be noted: -

- Half of the projects have timber production and erosion control as main objectives.
- 20% of the projects have objectives for planting multipurpose species.
- 10% of the projects have objectives for erosion control and fuel wood production.
- The remaining 20% of the projects have objectives for timber production only.

The choice of species to be used for afforestation involves extrapolation of information from elsewhere. That in turn, is determined by an actual data collected or through observation on a given site(s).

3.6 SPECIES SELECTION RESEARCH AND FRC

Forestry Research Center (formerly called Forestry Research Institute) was set up and inaugurated in 1961, through the financial support of GTZ (German Agency for Technical Cooperation) and US point four. FRC was also unsuccessfully incorporated into the Addis Ababa university science faculty, in 1967.

In 1975, UNDP/FAO started an initiative to support afforestation projects in the north mainly directed to the FRC seed section and species experimentations in various sites. These species selection sites are mostly found in south and southwest parts of Ethiopia (Husnia, 1997).

FRC has objectives of (EARO, 1997): -

- Providing forestry practice information necessary for policy and decision making at all levels by disseminating research output to extension worker, farmer or forest manager that leads to national agricultural development.
- Conservation, development and rational utilization of the nation's renewable resources such as forestry, soil, water and wildlife.
- Improving the production and productivity of woody plants.

Generally, research activities in the FRC have been undertaken with the objective of alleviating the priority problems in the specific subject area.

The objectives of FRC are fulfilled by focusing on the functional aspects, i.e., using a two way communication between users (e.g. farmers) and the researcher through which research needs emanate from decision makers and forestry work, and information gained through research is organized in a way suitable for decision-making (EPA, 1997; Berhanu, 1997).

Sections organized in FRC for the above purposes are (See Appendix I for FRC organizational structure) (Amare, Abate And Gebre, 1990; FRC, 1986; EARO, 1997): -

- Silvicultural section -the field experiments are mostly done in this section. Some of the field experiments are species selection, establishment of planting techniques and regeneration trials.
- Social- (agro-) forestry section –this section is designed to solve social needs and practices of tree species in and around their farm areas. Some of the experiments are socio-economic study and effect and acceptability of species (e.g. shade and shelter effect of different species).

- Wood utilization research center – focuses in providing information on the right utilization of woody species for timber uses.
- Forest management section- focuses on two studies, i.e., inventory of tree yield and tree measurement.
- Forest products and marketing section - engaged in surveying on the forest products (e.g. resins and gums) and their marketing conditions.
- Seed procurement and genetics section –engaged in supplying adequate and suitable seed for Ethiopian qualitative and quantitative needs, for most sites and uses.

The species selection task is done in the first three sections mentioned above. The studies undertaken by these sections, though, are used in the other sections namely, Forest management, Seed procurement and genetics, and Forest products and marketing. So, in an effort to incorporate the supplementary issues dealt with these sections after the determination of species to be selected in the area, the prototype also displays the information on seed and management of tree seedlings.

The species selection task in FRC, which is dealt in this study, involves comparing performance of large number of different species on one or number of sites and to select a smaller number for more intensive experiments (FRC, 1986). To accomplish the task, FRC uses experimental sites distributed presently in 21 locations in southern, northwestern and southwestern regions of the country (EARO, 1997). The research sections in FRC, which carry out research in species selection, are described below along with the factors each section considered necessary for the particular research undertaking.

3.6.1 Timber Properties Research

The timber properties are studied in the Wood Utilization Research Center (WURC), which is part of FRC. WURC studies mainly the mechanical physical and working properties of wood. These properties are indicators of usage of wood products properly and effectively. 35 species have so far been studied in the laboratory environment using observation and use of tester machines (for mechanical properties only). The species come from selected sites. The researchers determine the properties and give recommendations for the species displaying the best properties for the particular use.

The factors they work out in recommending the species are the following (WURC, 1985):

- ***The mechanical properties*** to be defined in this research area are the ability/strength to bend freely and regain normal shape, the ability to resist bending, relation between stress and strain, tensile stress to elongate it or the shearing stress tending to cause one portion to slide over. ***The chemical properties*** to be identified include effect of chemicals and natural defects. Chemical properties research is far too scanty, and cannot be applied in this research.

These two properties determine the species selection for poles, transport roads, building construction and et cetera.

- ***Wood working properties*** include machining properties, such as sawing, planning, finishing, and use of fasteners like screws and nails. Such information is used for determining whether the species can be of use to furniture and building purposes.

These properties in wood research are used as a basis in determining the benefit the species could give as a timber species. The information on these properties is given in the description part for every species. The reason for not including these properties in the rule is in the assumption that the user is not asked about these properties but what sort of benefit he wants. Then the expert simply gives the species and describes to the user that the wood displays such and such properties.

3.6.2. The Silvicultural Research

The silvicultural research screens those species adaptable and fast growing in a particular area. According to FRC (1986), the researchers use in their experiments site conditions, mainly rainfall, along with the soil conditions of the area, which are given as factors for adaptability. Rainfall is given emphasis because it is used as a basis to divide the country into *Bio-Climatic Zones*. The soil conditions they consider, among the various measures of soil, are the depth, texture, and PH.

Observations are made also about the survival rate (as another factor for adaptability). Here the researcher observes the rate at which a species survives, and decides, having at hand the objective data, whether a species survives in the area or not. The observations and decisions from the expert are needed because the data collected is not complete, and ideally, complete data should consist of all environmental and biological factors into consideration. Additionally, the longer (for decades) the research is conducted the more reliable the species selection would be. So, the researcher closely watching the performance of the species decides based on his experience, plus the data on hand.

3.6.3. Social Forestry Research

Within the Agro forestry research, the chosen species for a particular area are mainly studied to investigate their extent of acceptance and preferability by a particular community. This involves largely observation (EARO, 1997).

Species elimination trials in the section as well as observations in a specific area are also made for the purpose of selecting multipurpose trees (MPT) which fulfill suitable characteristics as erosion and soil loss control, adaptability to the sites and the like. These suitable characters are again matched with the species for a suitable purpose in the area.

Some experiments on the practice of Agro forestry are done, especially on alley cropping, i.e., planting a tree in a cropland. Most of the studies made are observational. These studies include soil and water conservation practices like erosion control, mulching ability, n-fixation, soil fertility increment, shading and sheltering, and water balancing. The other studies include the social benefits of tree planting in relation to marketability, like income generation, increment of yield when planted with crops. Still, some studies focus on the benefit the tree residues or wastes have for animals and bees as fodder (food). Such studies are made in collaboration with the International Livestock Research Center (ILCA).

The role IT plays in supporting research activities at FRC is minimal since, IT staff and equipments hardly exist in the organization. Only the librarian and his assistant have formal training in IT to manipulate the library catalogue database. The researchers are assisted by librarians in searching a document in the database. There are also limited numbers of computers in the organization; only 14 computers are available in the center. Except the two found in the library, all the other computers are used for word-processing and spreadsheet services only.

But recently, changes are being observed in the FRC. Additional computers for the library and the offices were being purchased and trainings for all members of the center in MS-word, MS-access and MS-excel have been given. Since plans are underway in training researchers and other supporting staff to train in other statistical software's (SAS software, for instance) and systems applications. The already given training paves the way to the researchers and other support staff in successfully participating in planned applications next time.

So expert system is chosen with the idea that it could solve problems in the center, and those researchers have the potential curiosity to be trained in computer applications like expert systems.

Therefore, concepts taken from the study and long time experience of experts from the above described research sections in species selection are taken and put into a hierarchical manner, in consultation with experts in the area.

CHAPTER FOUR
KNOWLEDGE ACQUISITION

4.1 INTRODUCTION

Knowledge acquisition refers to any technique by which computer systems get the knowledge they need to perform their tasks. In short, it is a process of extracting and structuring knowledge in the particular domain/ field by the knowledge engineer (Wormell, 1987).

Though knowledge acquisition process is most important, it is also tedious and least formalized system of the expert systems development activities. As a result, it is considered as an art and a bottleneck in developing expert systems (Masuch, 1990).

This is because: -

- There are no enough knowledge acquisition tools available.
- There has been little success so far with domain shells, which are designed to allow domain experts to construct their own expert systems.
- Domain knowledge is hard to extract from domain-knowledgeable people.

This chapter discusses the process of knowledge elicitation and modeling, and the resulting concepts acquired about the species selection task in FRC.

4.2 STEPS IN KNOWLEDGE ACQUISITION

Two knowledge acquisition techniques have been applied in this study, namely, knowledge elicitation and knowledge structuring. The former deals with getting knowledge from the expert, while the later suggests the process of modeling the concepts involved in the species selection process. The two techniques are described briefly below.

4.2.1. Knowledge Elicitation

Knowledge elicitation, the most important branch of knowledge acquisition, signifies the extraction of knowledge from human expert(s) to build an expert system. The knowledge elicitation task involves finding at least one expert in the domain who's willing to, has time and is able to, provide his knowledge (Benchimol, Levine and Pomerol, 1987).

For this reason, three experts in the three departments have been interviewed, i.e., one from wood properties, one from social forestry and one from silvicultural sections. The experts have 15-30 years of experience working in their respective departments. Three junior researchers are also used in times of absence of the experts, and to clarify definitions and ideas found in the handbooks and textbooks.

Non-structured type interviews were held with experts toward the beginning of the interview sessions, for the purpose of familiarizing with the expert researcher and the subject matter of species selection. Afterwards, a series of structured interviews have been conducted with an objective of getting the concepts, variables and the resulting ranking of species according to the priority they assign to these species, of their adaptability or usability.

Documents of FRC have also been revised to support the interview sessions. Some of the reasons in reviewing documents are: -

- To remind the expert ideas and concepts that he might have forgotten in the previous interview, so that the expert explains ideas better.
- To build a species description, in addition to what is communicated from the expert.
- To have a clearer understanding of the theories and concepts involved in species selection, thereby, enhancing the communication between the developer and experts.

4.2.2. Knowledge Structuring

Knowledge structuring is a process where the knowledge engineer uses concepts discovered during the knowledge acquisition sessions to build a model/ representation of the expertise. From then on, the model 's development proceeds with stepwise refinement (Parsaye and Chignell, 1988).

The representation of concepts of the species selection is done using hierarchical tree structure. In a hierarchical tree structure, a goal, in this case, species selection is put at the top, or highest level of hierarchy. According to Benchimol, Levine and Pomerol (1987), this is because a goal is dependent on key factors, and it is only dispensed when one or more of key factors have been inferred. The primary factors are put immediately below the goal, and below them, secondary factors and so on.

Such reasoning approach in knowledge acquisition and representation is called top-down or deductive approach, because a person starts from general and overall concepts, gradually leading to elicit details of the topic. The reverse approach, the inductive approach starts from specific cases and reaches at a general concept (Ebrahim, 1999).

As shown in figure 4.1, an ellipse/node, the hierarchy, represents the concepts, whereas, direction of the arrows moves from the higher to lower level. A node on the graph is linked to another node from the neighborhood by the path/arrow. The depth/level of the root (when the node is further classified) increases every time nodes are added. The node with no successor is called a Hanging Node or a Leaf. Such organization of concepts creates categories that contain concepts with similar properties. It is the way experts awaken their long-term memory and explain the facts to their users (Richardson, 1995).

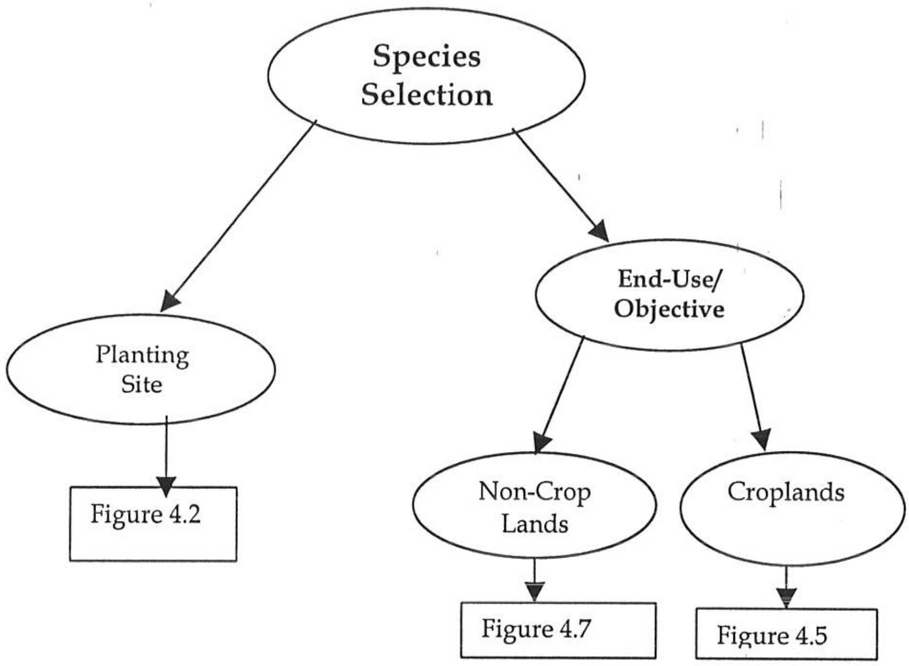


FIGURE 4.1. CONCEPTS IN SPECIES SELECTION

So, the expert follows the main concept depicted in figure 4.1, where each expert expands the concepts in the nodes by branching them whenever necessary.

4.3 CONCEPTUAL MODELLING

In modeling the prototype expert system, sample species of frequently researched species have been selected both in the category of widely planted and underutilized species as shown below (The chosen species are listed in table 4.1 below).

- Category of widely planted species- According to EFAP (1994) eucalyptus species comprise about 60% of state plantations, while 30% comprise Cupressus and Pinus species. Because the Ethiopian rural community knows these species better, and improved understanding of their use leads to proper species utilization.

- Category of underutilized species - 10% of all plantations comprise other species' plantations like Juniperus and Olea, occupying most of the woodlands and shrub lands. According to EARO (1997), species like Acacia can be of great help in replacing the over utilized species, like eucalyptus, in the farming community.

The Species chosen are mainly indigenous (e.g. Hagenia, Cordia and Juniperus). But some exotic species like Pines, Cupressus and Eucalyptus are also included in the list because these species are known and planted widely.

Acacia albida	Casuarina equistifolia	Gravelia robusta
Acacia saligna	Cordia abyssinica	Hagenia abyssinica
Acacia Senegal	Croton macrostachys	Juniperus procera
Acacia tortillis	Cupressus lusitanica	Olea carpensis
Albizzia lebbek	Eucalyptus camaldulensis	Pinus radiata
Azandrichta indica	Eucalyptus saligna	Pinus patula
	Eucalyptus globules	

Table 4.1 the sample species used in prototyping the expert system.

The species in table 4.1 have not been listed for every concept in the model of species selection represented below from figures 4.1-4.12 by the hierarchical method. This is because, as can be observed in the sample rules found in Table 4.2, the sites (which include zones and soils) determine the species to be listed. And these tree species are again narrowed down when it comes to the benefit (end-use) rules.

Species	Sites											
	1	2	3	4	5	6	7	8	9	10	11	12
Acacia albida	0.6				0.5	0.4						
Acacia saligna				0.9				0.8	0.5	0.8	0.6	0.7
Acacia Senegal						0.9						
Acacia tortillas		0.5				0.7						
Albizzia lebbek				0.4	0.4							
Azandrichta indica				0.5	0.8							
Casuarina equistifolia		0.7	0.6									
Cordia abyssinica			0.9				0.8	0.9				
Croton macrostachys							0.8			0.5	0.7	0.7
Cupressus lusitanica	0.8	0.7	0.3	0.5				0.3	0.9		0.7	0.8
Eucalyptus camaldulensis	0.4	0.6	0.6	0.9				0.8	0.5	0.9	0.5	0.7
Eucalyptus globules		0.8	0.8				0.7	0.5		0.7	0.7	0.7
Eucalyptus saligna							0.8	0.7				
Gravelia robusta	0.5	0.9	0.4			0.5	0.9	0.4				0.7
Hagenia abyssinica						0.8	0.8					
Juniperus procera	0.4				0.6		0.9		0.9		0.8	0.8
Olea carpensis	0.3	0.3	0.3							0.6		
Pinus patula							0.6		0.7		0.6	0.4
Pinus radiata	0.6	0.7							0.8			

Table 4.2 the certainty factors of the sample species with sites

KEYS TO THE SITE NUMBERS

SITES 1 -Loam semi arid low lands
SITES 2- Sandy semi arid low lands
SITES 3 -Clayey semi arid low lands
SITES 4 - Loam dry lowlands
SITES 5 -Sandy dry lowlands
SITES 6 -Clayey dry lowlands
SITES 7-Loam semi wet land
SITES 8 -Sandy semi wet land
SITES 9 -Clayey semi wet land
SITES 10- Loam wet region
SITES 11 -Sandy wet region
SITES 12 -Clayey wet region

N.B. Those species which have not been assigned to a site signify that they are not adaptable to the area at all, or there is no enough information /experience in the part of an expert to conclude on a certain value of CF.

As shown in figure 4.1, the two steps to select a species are, first to decide on the planting site for the species, and second, to pick the end-use of the plantation activity. Site selection for the species is necessary because every site have unique character, where a species character cannot be predicted unless a research is conducted in the area. The planting objective is also different for every species. For instance, few species are vigorously increase the yield of the crop, still others affect crop yield negatively. So, by combining the above two entities, we get a species both adaptable and utilizable to the area.

Lastly, the model construction is done after iterative process. This is done by submitting three different drafts, one after the other, to experts. The iterative process is better understood in the following descriptions of things considered as priorities in structuring the concepts in each draft.

The first draft is prepared after the preliminary interview and choice of documents to be used for research. So, the draft prepared was more theoretical than practical. As the expert revise the draft, ideas on how the expert further formulate his concepts and the reason for doing so is discussed. This has taken lots of discussion sessions.

The second draft, is out, based on the discussions from experts. In the revision of the second draft, the developer asks the experts to consider factors based on the need of the farmer, research availability and the Q&A sessions with the user, to help change the model to production rules. The discussion of the experts and revising the draft together brings about the final draft of the model.

4.3.1 Planting Site

The site conditions affecting the planting site are soil, temperature, altitude and rainfall. According to the experts, due to the simplicity and consistency needed for research works, rainfall and soil are the only ones to be used in all researches. This is one of the many rule of thumbs used by experts to easily explain the entities used (e.g. Planting site) in their research.

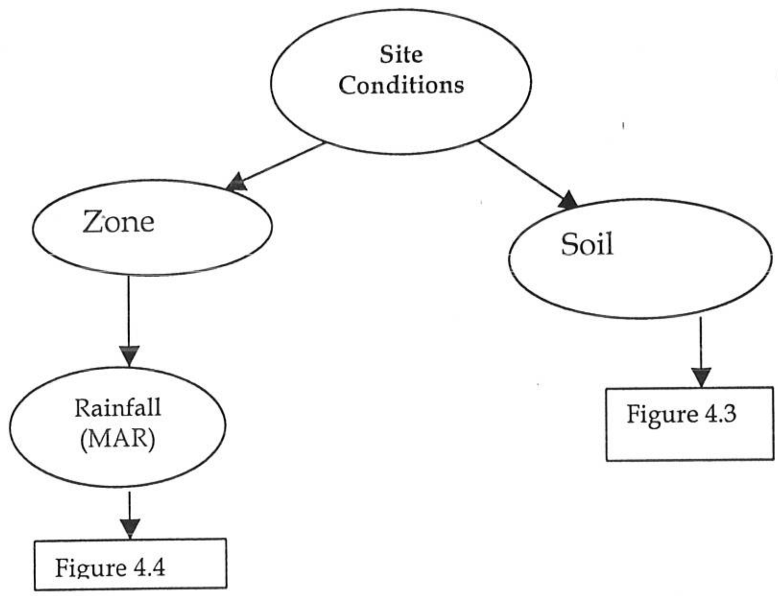


FIGURE 4.2. CONCEPTS IN PLANTING SITE

The Soil - the procedure followed by experts in site selection is classifying the site to broad climatic conditions as shown in figure 4.2. The site conditions narrowed down to include the local (particular site) conditions. Soil PH, texture and depth are the factors used by the experts in determining the local conditions (see Figure 4.3).

But it is only the soil texture that is used frequently, but other soil characters are used within the rules to differentiate between concepts. The soil structure is frequently used mainly because it is easier for the user to tell the texture of the soil (whether its sand, clay or loam.) than describing concepts like shallow or saline soil. The latter descriptions are more experimental, necessitating a soil scientist rather than a forester or extension worker. So, the PH and depth of the soil is presented here as concepts merely used by the expert for documentation rather than communication with the user.

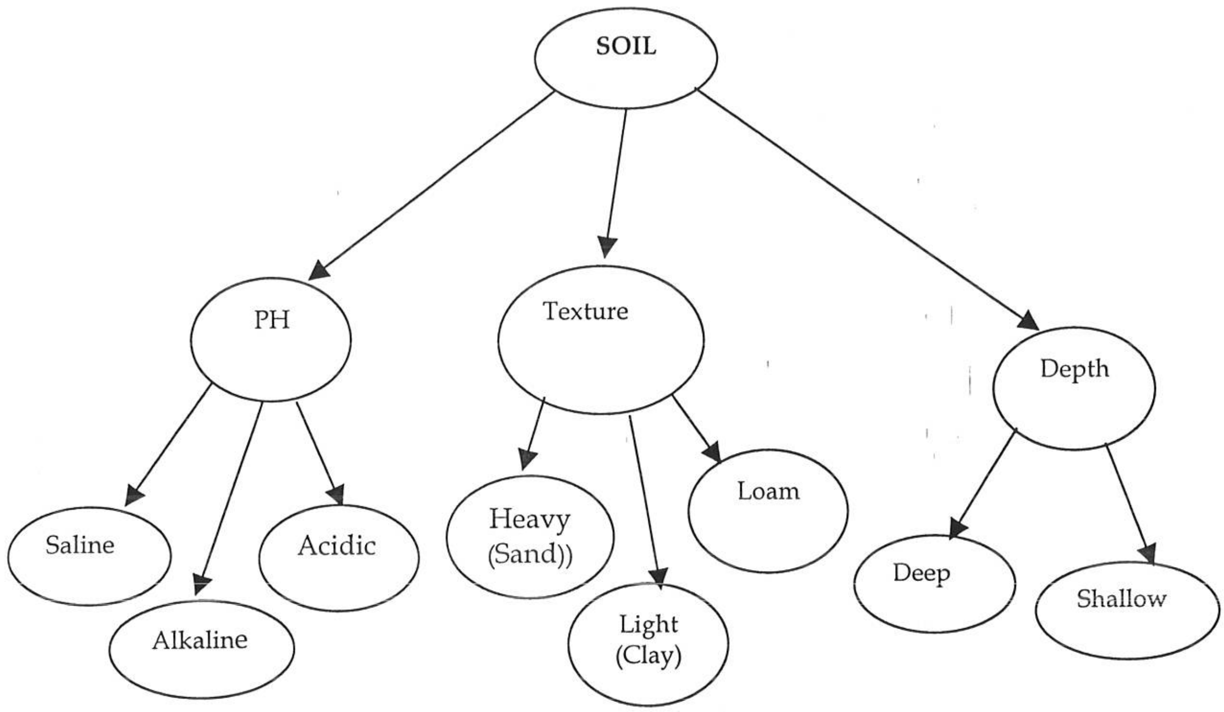


FIGURE 4.3. CONCEPTS IN SOIL DETERMINATION

Mean annual rainfall (MAR)- is the measure of the overall rainfall or precipitation in the area. For the purpose of research site categorization, the MAR has been divided by the research center (FRC) into five major bio-climatic zones as shown in figure 4.4. In particular, zones are divided into:

- Semi-desert: mean annual rainfall ranging 0-350mm.
- Dry lowlands and plateau: mean annual rainfall ranging 350-700 mm
- Semi-arid lowlands and plateau: mean annual rainfall ranging 700-1050mm
- Semi-wetlands (and plateau): mean annual rainfall ranging 1050-1400mm
- Wet land: mean annual rainfall more than 1400mm

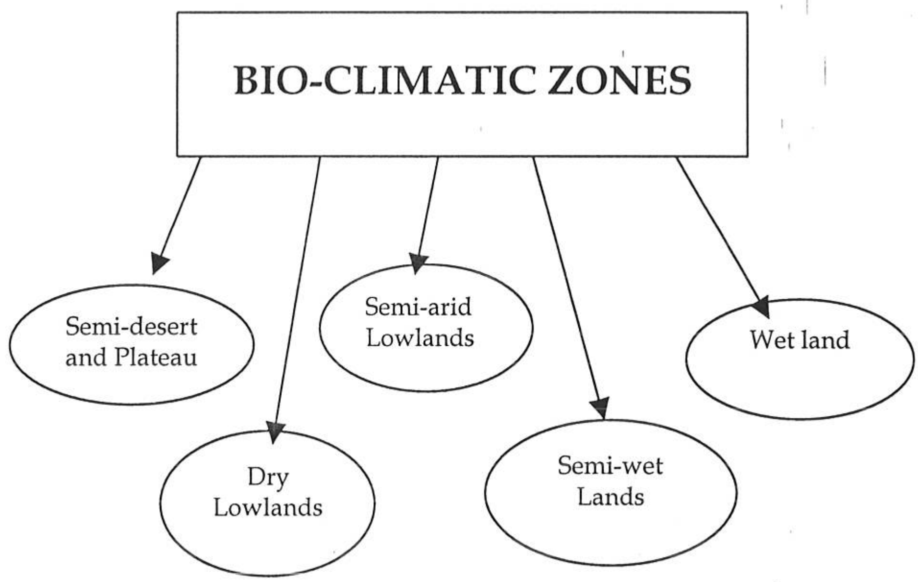


FIGURE 4.4. CONCEPTS IN BIO-CLIMATIC ZONES

4.3.2 End-Uses of Trees / Shrubs

Tree species contribute a wide range of goods and services to the rural community. Trees may provide products and services like food, shelter, energy, cash income, and improvement of soil fertility for crop production and yield, crop pest control and others. In addition, such practices are intended to protect and improve quality of natural resources, like soil and water.

a) Tree planting in cropland

Tree planting in croplands has many uses. Among the uses, which are identified by the experts, includes production increment of the surrounding crops, through nutrient cycling and provision of soil moisture in the form of litter/mulching. As shown in figure 4.5, trees provide windbreak, shade and shelter, and yield increment to the crop. If the landscape is sloped, a species with the purpose of stabilizing the land, soil conservation activities like terracing, contour strips and small structures from stones and grasses are used.

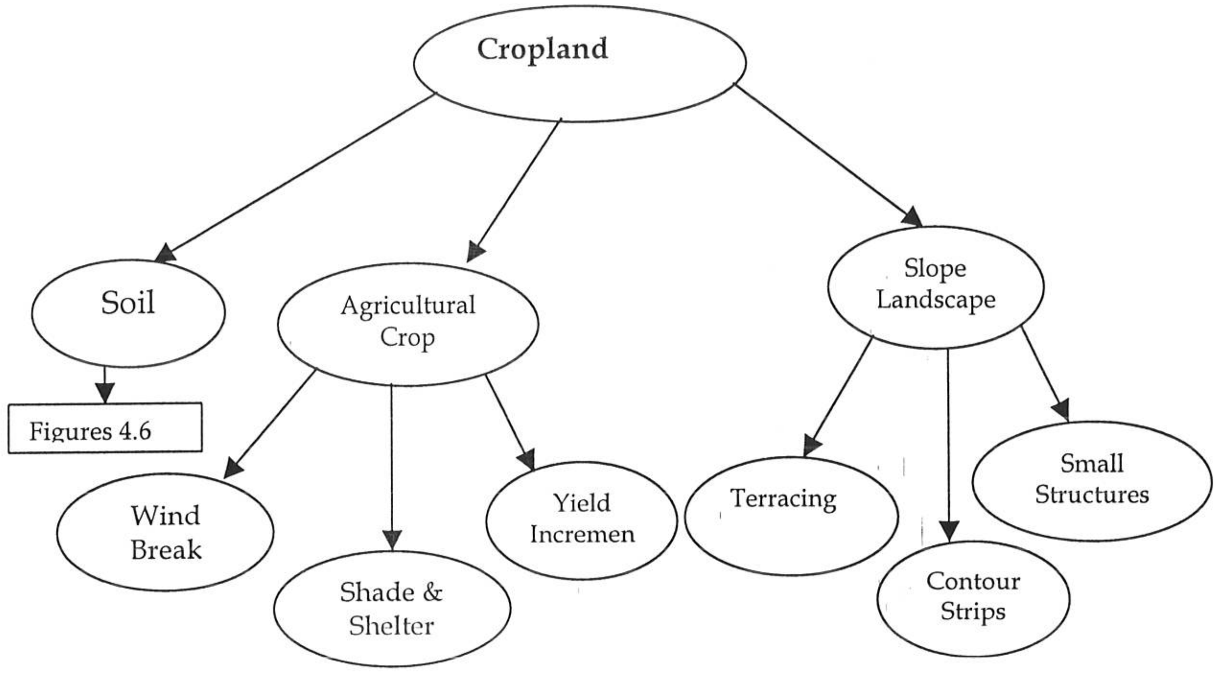


FIGURE 4.5. CONCEPTS IN TREE USES IN CROPLAND

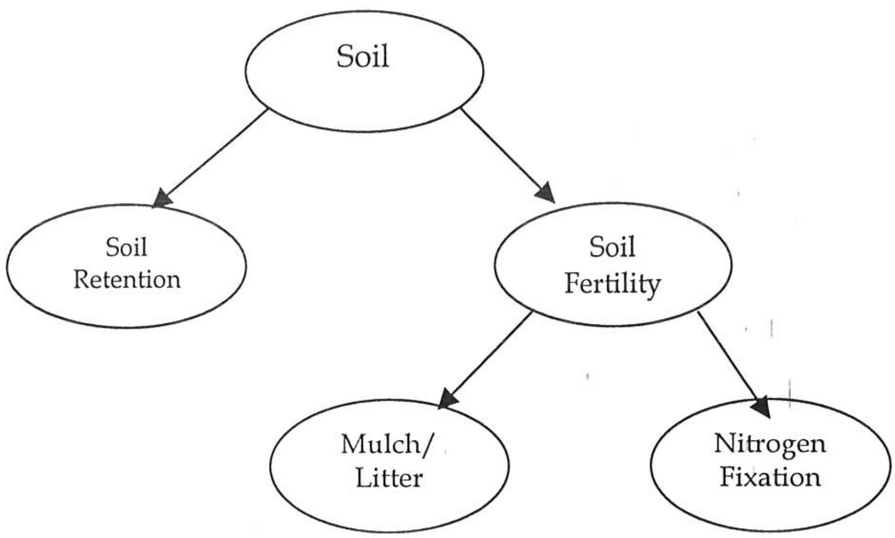


FIGURE 4.6. CONCEPTS IN USES FOR SOIL FERTILITY AND RETENTION

b) Tree planting in non-cropland areas

Non-croplands are areas where agricultural crop (Sorghum, Millet, Teff etc) cultivations are not practiced. This is simply because such areas are waste lands, around homes, places for collecting firewood and timber (woodlots) or grazing (pasture) lands for animals.

- **Wastelands** - are landforms not suitable for crop cultivation. They are classified generally as gullies and swamps. So, to make them usable, appropriate species are planted (see Figure 4.7).

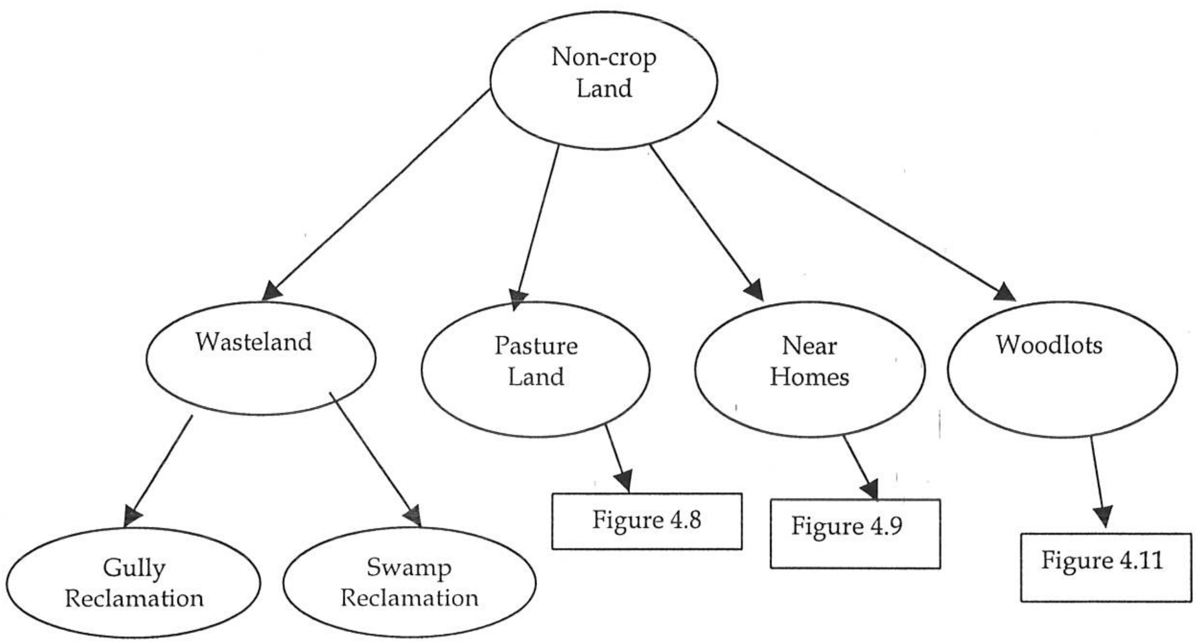


FIGURE 4.7. CONCEPTS IN NON-CROP LANDS.

- **Pasturelands** - are primarily used for grazing, but resistant species on the area can be planted, whether these species resist draught, poor drainage or pest (See Figure 4.8).

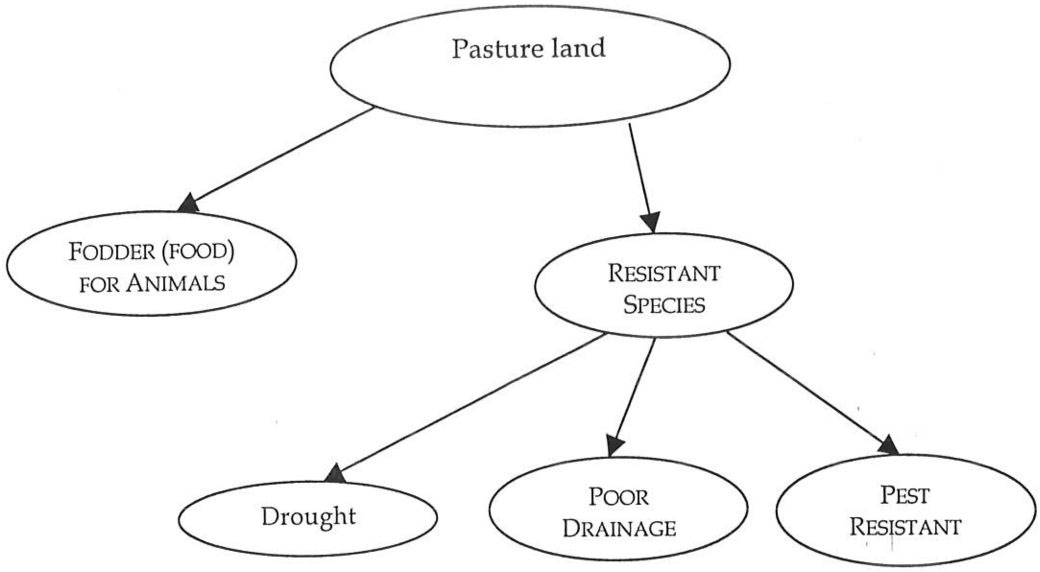


FIGURE 4.8. CONCEPTS IN TREE USES IN PASTURE LAND.

- **Places near homes** - are preferable spots of plantations, especially for household women. This is because women could plant trees in their backyards to avoid traveling longer distances to markets to buy simpler food items like fruits and spices. The tree species can also be used as sources of income (as shown in figures 4.9. and 4.10.).

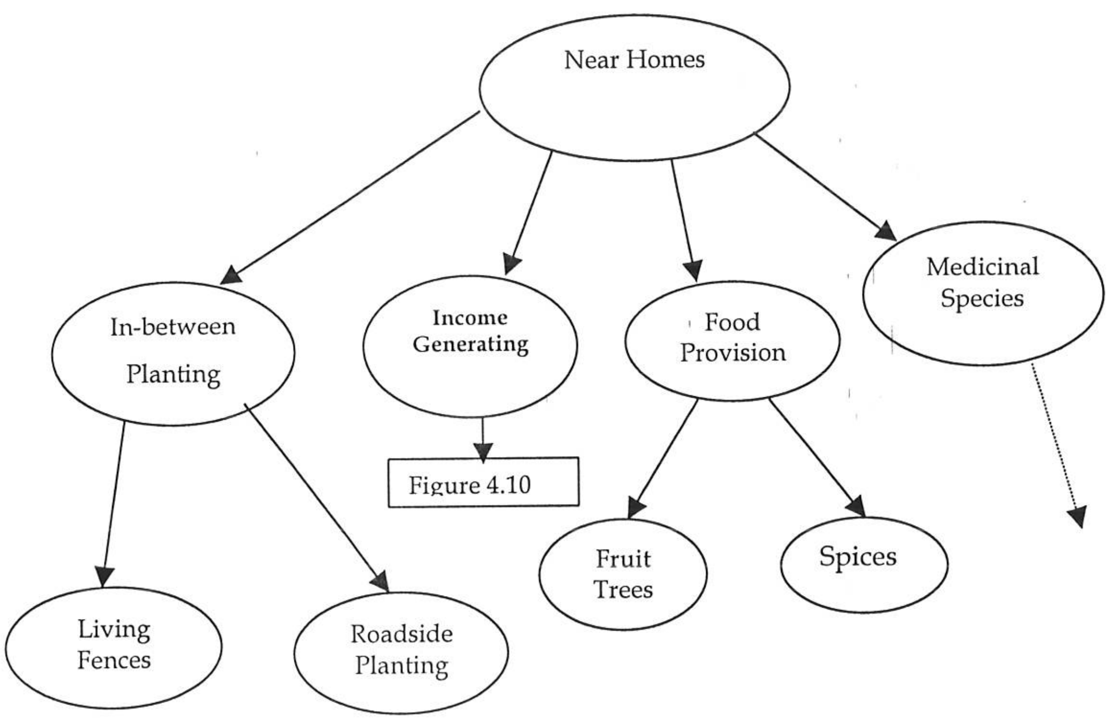


FIGURE 4.9. CONCEPTS IN TREE USES NEAR HOMES

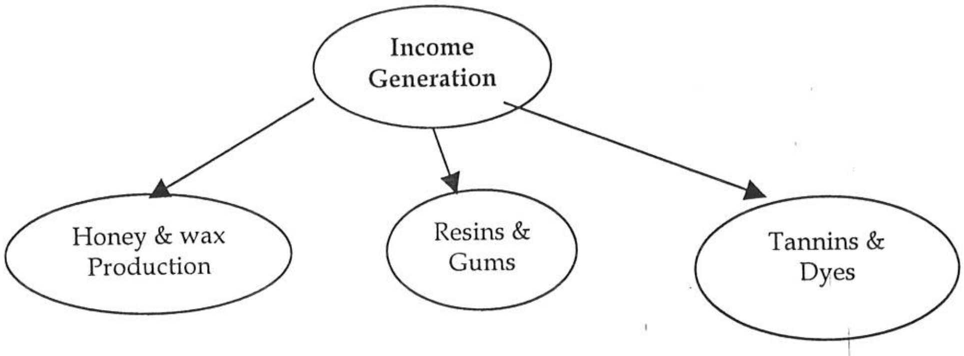


FIGURE 4.10. CONCEPTS IN INCOME GENERATION.

- **Woodlots** - Tree cutting for search of wood products is becoming the main cause of deforestation. The farmer used tree species for fuel wood and charcoal (Figure 4.11), as the main rural energy source while timber is used for light construction, furniture, poles and posts, and agricultural tools (see Figure 4.12). Expansion of woodlot plantations of wood and energy producing trees has been considered to be one way of slowing down the rate of deforestation. Due to these reasons, emphasis is given in this study to the needs of farmers and extension workers when drawing out concepts in timber production.

The following figures (figure 4.11 and 4.12) show hierarchically the uses of the woodlots, which are also the basis for species selection:

CHAPTER FIVE
KNOWLEDGE REPRESENTATION

5.1 INTRODUCTION

So far, the domain knowledge has been modeled using the hierarchical structure. Using the structure, the knowledge is represented into a knowledge base. In this chapter, the knowledge representation technique, the rule based system and the Certainty Factor (CF), along with the development and integration aspects of SPEX, the prototype species selection expert system, has been discussed.

The representation technique used in this study, rule-Based representation is a popular knowledge representation method in expert systems. A rule consists of 'IF' and 'THEN' part. If the IF part of the rule is satisfied, the THEN part (its problem solving action) can be concluded (see Appendix II for rules in knowledge base of SPEX).

The popularity of Rule based representation is because rules (Stefik, 1998;Ebrahim, 1999): -

- Take more general nature of human reasoning.
- They are independently represented, i.e. they are easy to add, delete and modify, the knowledge base.
- They are easy to construct; most importantly can be linked so that the conclusion of one rule can form the hypothesis of another rule.

Because of the reasons mentioned above, the current undertaking considered rules as knowledge representation methods. So, using an IF -THEN rule-based system, the Prototype Species Selection Expert System (SPEX) is developed.

5.2 KNOWLEDGE BASE

The knowledge base is a collection of database of knowledge, which, in this case, is designed using rules. The rule consists of statements connected by 'If' (hypothesis/premise clause) and 'Then' (the action or concluding results).

The following sample rules show how the rules are constructed in the knowledge base of the expert system: -

topic 'recommended species'.
set_number_of_values ('recommended species',1).

SITE RULES
Rule 1
if ?rainfall is 701-1050 mm
and ? soil texture is clay
then site is 'clayey semi-wetlands'

Rule 2
if ?rainfall is 351-700 mm
and ? 'soil texture' is sand
then site is 'sandy dry lowland'

:
:

END-USE RULES
topic 'recommended species'.
set_number_of_values ('recommended species',1).

Rule I
if ?site is 'sandy dry lowland'
and ?'land use' is 'crop land'
and ?'crop land' is soil
and ?soil is 'soil fertility'
and ?'soil fertility' is 'nitrogen fixation'
then 'recommended species' is 'Juniperus procera(0.4)Acacia albida(0.4)'.

Rule II
if ?site is 'sandy dry lowland'
and ?'land use' is 'crop land'
and ?'crop land' is soil
and ? soil is 'soil fertility'
and ?'soil fertility' is mulching
then 'recommended species' is 'Juniperus procera(0.4)Albizzia lebbek(0.4) Acacia albida(0.3) Azandrichta indica(0.24) '.

```
⋮
RULE XI
   if ?site is 'clayey semi-arid lowland'
   and ?'land use' is 'crop land'
   and ?'crop land' is soil
   and ?soil is 'soil fertility'
   and ?'soil fertility' is 'nitrogen fixation'
   then 'recommended species' is 'Eucalyptus camaldulensis(0.5)Eucalyptus globulus(0.5)'.

Rule XII
   if ?site is 'clayey semi-arid lowland'
   and ?'land use' is 'crop land'
   and ?'crop land' is soil
   and ? soil is 'soil fertility'
   and ?'soil fertility' is mulching
   then 'recommended species' is 'Cordia abyssinica(0.7) Olea carpensis(0.3) '.
     :
     :
    if ?site is 'loam dry lowland'
  and ?'land use' is 'non crop land'
  and ?'Non-cropland' is 'near homes'
  and ?'near homes' is 'in-between planting'
  and ? 'in-between planting' is 'roadside planting'
  then 'recommended species' is 'Cupressus lustanica (0.6) Eucalyptus camaldulensis(0.4)Pinus
   radiata(0.2),
    ⋮
   if ?site is 'sandy semi-wet land'
   and ?'land use' is 'crop land'
   and ?'crop land' is soil
   and ? soil is 'soil retension'
   then 'recommended species' is 'Pinus radiata(0.6), Juniperus procera(0.5) '.
```

The numbered rules (rules 1and 2) above are the site rules. These rules are linked with usage rules I, II, III.... by "Site" statements (i.e. phrases like "? Site is 'clayey semi-arid lowland") which are found as conclusions in the former rules, and as hypotheses in the later rules.

The '?' (Question mark) indicated in the rule statements show, the value of the entity. For instance, "? Soil texture is clay," indicates that the value of soil texture among other attributes of soil texture is clay. Other attributes or values of soil texture are loam and sand (See Figure 4.3 for concepts of soils).

On the other hand, the word "is", in every rule statements, equates the left hand and right hand side values of the statement in the rule. For instance, in the previous example, the value (?) of "soil texture" is equated with "clay". This means, whenever soil texture is called in the rule clay is implicated. As a result, the knowledge- pro expert system shell takes "=" (the equal sign) and "is", as same operators.

Because the number of values is set to be one, more than one word is put in a single quotation.

For instance, topic 'recommended species'.
set_number_of_values ('recommended species', 1)..

The main topic, 'recommended species', and its rules are bound to enter only one value, using the declaration of set_number_of_values in the beginning of the sentence. The rules are then linked together and effectively searched by the inference engine. Sample rules in the knowledge base, representing one site and species according to their usage are presented in table 4.2.

5.3 THE EXPLANATION FACILITY

The explanation facility of SPEX is endowed with a human like trait of explaining its line of reasoning to offset suspicion and increase acceptance. This is especially necessary when the system is handling uncertainty.

The elucidation facilities incorporated into SPEX are the following:

- **Why facility**: - gives definitions to clear the seemingly obscured ideas or concepts in the forestry domain language. For instance, in an interface where the user is asked.

What sort of land use system do you have in your area?

Cropland

Non-Cropland

The user might want to know exactly what the exact definitions these alternatives have. By clicking the button named WHY, the user gets the following explanation.

Non-croplands - are areas where agricultural crop cultivations are not practiced. In other words, these areas could be wastelands, around homes, places for collecting firewood and timber (woodlots), or pasture lands.

Croplands - are areas where agricultural crop (Sorghum, Millet, Teff etc) cultivations are practiced. These areas are also called farmlands.

- How facility: - answers how the rules come up to conclude with a specified Certainty Factor. In short, the facility answers how the system reached to the specific conclusion and with that Certainty Factor.

When the recommendation at the end of the Q&A session displays

THE FOLLOWING SPECIES ARE STRONGLY ADVISED TO BE PLANTED FOR THIS PARTICULAR PURPOSE:-
Acacia mearnsii (0.6)'.

And if the user clicks the button, 'How', an explanation is displayed

The calculation is done to ~~Acacia mearnsii~~ of the two CFs of the species by multiplying the site CF, which is 0,8 with the usage CF of 0.8 . And the result is 0.64 $\approx$ 0.6.

- Using the hypertext system supported by knowledge pro, further explanations are given, in the 'INDEX' part of the user interface.

5.4 INFERENCE ENGINE

As with other expert systems, the inference engine in the Prototype Species Selection Expert System, SPEX, has the task of applying the domain knowledge in the knowledge base using a search strategy of backward chaining that facilitates the probing back to the final goal, the recommended species.

The inference engine is part of and the built-in code of the knowledge pro expert shell. The shell offers minimal direct control over the inference engine to the programming, because it helps the user put only the domain knowledge into the system, and leave the inferencing (the search of knowledge) to the inference engine. But the syntax used in directing the inference engine to search for the goal behaves in the following manner:

```
topic 'pasture land'.
Ask ('what do you want to do in your pasture land?', 'pasture land',
['fodder for animals', 'non-timber products']).
end.
```

This syntax is used to search the intermediate goal of the system, pasture land. To get to the 'pasture land', the system asks the user to enter one of the two attributes of 'pasture land', as shown above.

To get to the ultimate goal, the same syntax is used, but this time, using the feature found in knowledge-pro ,text:

```
Topic 'recommended species'
text ('the following species are recommended strongly:', ?'recommended species').
End.
```

The inference engine then searches for the values of (as '?' means) 'recommended species' and then lists in the end, as a recommendation. Here is an example:-

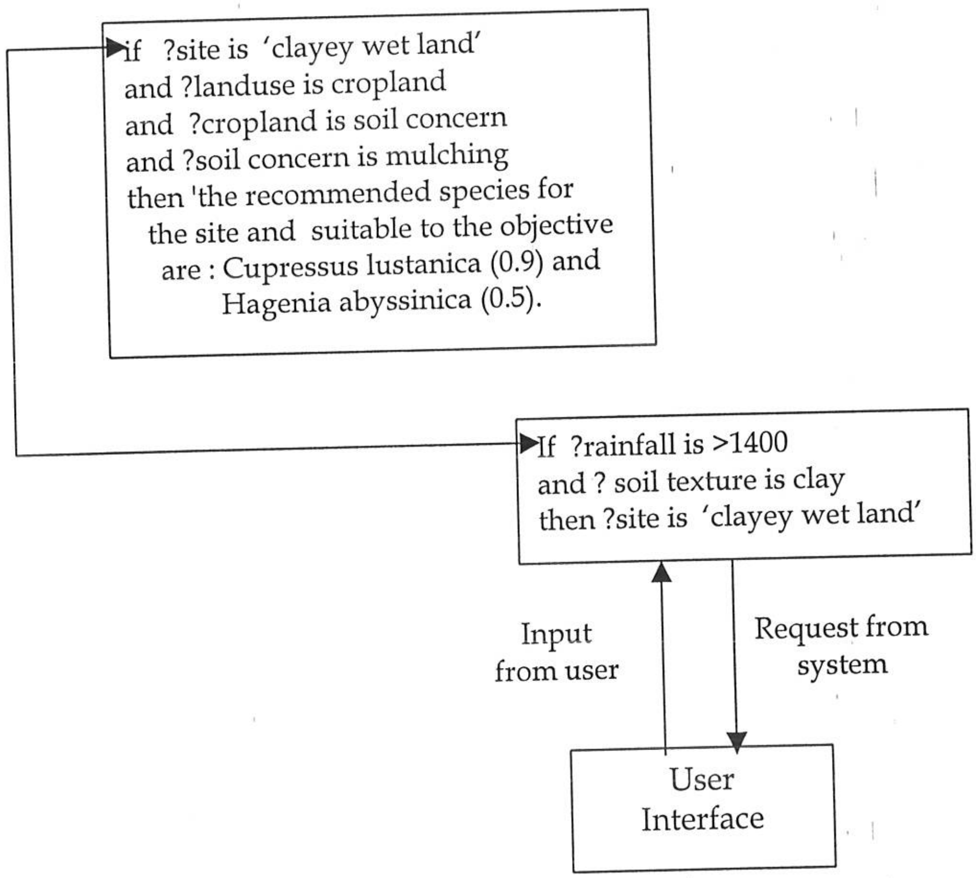


FIGURE 5.1 BACKWARD CHAINING IN SPEX (THE PROTOTYPE EXPERT SYSTEM.)

The rules in the knowledge base are put in three levels, where the inference engine as a backward chaining strategy, first searches for the value of site. But having no information, the engine searches backward to get that zone which has some value. But again the search moves backward and ascertains that the value of zone is decided by the value of rainfall, or specifically, the mean annual rainfall (MAR). Once more the inference engine moves backwards to ask the user to input a number as value of rainfall. Figure 5.1 shows the backward engine strategy.

Subsequent questions by the system therefore starts from rainfall, and then moves to soil texture, land use and so on, until it infers and matches user selection with the rules in the knowledge base, so that the system gets to the goal and recommend species.

5.5 ASSIGNMENT OF CERTAINTY FACTORS

a) The ranking for uncertainty assessment

The human expert gives a rank of the best performing species in a particular area or objective, running from the most uncertain at one end to the least uncertain at the other. The ranking is then changed to the Certainty Factor to signify the belief/confidence of the human expert on the success of the species through the discriminating mechanism.

Generally, all ranking methods can be classified in terms of a single or combinations of the following basic ranking techniques (Parsaye and Chignell, 1988):

- **Ranking by selection** -selection techniques pick out the objects in order, using the general rationale of " select the next smallest" until the list of objects is exhausted, and an ordered set of selections is obtained.

- **Ranking by insertion**- insertion techniques take the i^{th} object to be sorted, and inserted into the ordered list created from the previous insertions.

- **Ranking by exchange** -exchange methods successively compare pairs of objects in the list, by swapping or sorting the list for all N pairs of objects, until the entire list is ordered, as shown in Figure 5.2.

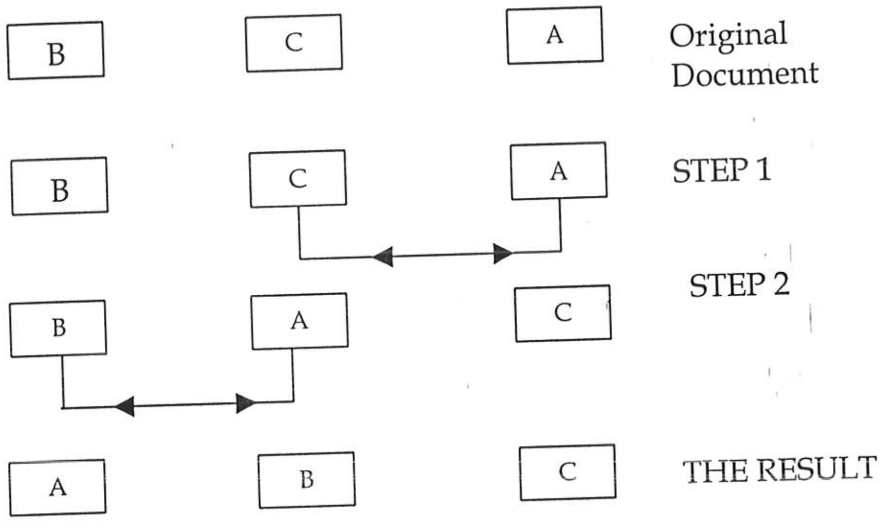


FIGURE 5.2 THE RANKING BY EXCHANGE

Of the three ranking methods, *Ranking By Exchange method* was used in this work. The ranking system is chosen in view of the limited list of species that is used in the study. As such, every pair of species is checked to each other by the human expert .The development of rules is also understood better, whenever the expert reasoned out why the particular species is better in performance than the other, taking pair of species at a time.

Generally, an expert assigns the ranking on the use of species, based on species character for a given benefit, and acceptability of the species in the community. The assignment of site CF for species is done based on researches and observational experiences on adaptability of species, as explained briefly in sections 3.6.2 and 3.6.3.

Then, as the experts rank (swap) the species, they assign values ranging from 1 to 10. Ten (10), in value, means best performing or highly acceptable species in use/adaptability to the area, while, 1 means poorly performing or with the least acceptance and adaptability to the area. Limited values (1-10 only) were given in the assumption that values exceeding 10 make the value assignment more bizarre and difficult for the expert. These values depend on the strength of belief by the expert that the species is more adaptable or useful for the given area or need when compared to the other species in the category. The values are simply changed to percentage by multiplying them by 10, which gives us the CF.

b) Combining or Updating Rules

In this prototype, two types of rules are assigned Certainty Factors (CF). The CF in one of the rule categories describe the belief the expert has in the adaptability of the species in the area (using the factors of rainfall and soils), but in the other rule categories, CF assign for the belief of the expert that the particular species is suitable for specific benefit

But the problem is, the conclusion should have only one Certainty Factor, the combined CF. So, to calculate the final/ combined confidence in the hypothesis, we need to combine the two CF provided by these rules.

The formula used in this study for combining the CF is: -

$$CF(C) = Y\ MAX\ (0, CF\ (B)) = Y*X.$$

The formula explains that to get the revised CF, the CF for B is multiplied by the existing CF for C. Details on the formula are discussed in section 2.5.2.3.

The following rules and their assignment are good examples to the above formula:

If ?rainfall is 701 -1050 mm
?'soil texture' is ?sand
Then site is 'sandy semi-dry lowland'
And the recommended species are the following:

SPECIES	SITE CF
Juniperus procera	0.9
Acacia melanoxylone	0.7
Eucalyptus saligna	0.8
E globulus	0.9
E. camaldulensis	0.7
E. grandis	0.7
Pinus radiata	0.7

Table 5.1 species selected based on site information.

The site conditions (soil and rainfall) limit the bulk of species available, to a few of them, ranked according to the degree of confidence of the expert to the plant for a particular objective. In table II, species, which are suitable for the site, are given by the expert a CF value. These species are used in this study, to calculate and combine with CF of the usage CFs that is explained in the next paragraph.

To move on to the next statement (s), The Certainty Factor, for the site conditions is combined and calculated with that of the CF of the species given to the soil fertility among the species already chosen for the suitability for the given site in the following manner.

If ?'land use' is cropland
and ?cropland is soil concern
and ?soil concern is yield increment
then 'the recommended species for the site and
suitable to the objective are:

Juniperus procera CF 0.96
Acacia. melanoxylone CF 0.91

This is how the combined CF formula shown above is calculated as follows: -

Site CF for the species =site.

Soil erosion CF for yield increment =yield.

So, the combined CF for each species = site * yield.

Then, the species will be put in descending order (on priority basis) of their combined CF.

SPECIES	SITE	YIELD	COMBINED CF
Juniperus procera	0.9	0.6	0.5
Acacia melanoxylone	0.7	0.6	0.4
E. saligna	0.8	0.2	0.2
Eucalyptus globules	0.9	0.2	0.2
E. camaldulensis	0.7	0.2	0.1
E. grandis	0.7	0.2	0.1
Pinus radiata	0.7	0.2	0.1

Table 5.2 the combined CF in species selection CF calculation.

Only two of the species with the combined CF of greater than 0.2 are considered. This is because the certainty or degree of belief of the expert is so low as 0.2 that it is better not to consider these species. In addition, it has been a tradition for most expert systems since the first user of CF (MYCIN) not to consider CFs less than or equal to 0.2 (Schmoldt and Martin et al, 1989; Stefik, 1998).

5.6 THE USER INTERFACE

The user interface is the interface found between the system and the user.

The user interface works in the following manner:

The interface is opened with the home page with a message "Welcome to the prototype expert system in species selection. This advisory system helps you in selecting a species adaptable and usable in your area" - that welcomes the user (as shown in figure 5.3).

All interface windows consist of the following buttons: -

- **Back and Quit buttons** - for returning one step forward or stopping the application, respectively.

- **Index button** - gives information on species management issues like:
  - Local naming of the species, in a widely spoken Amharic and Oromo languages.
  - Species description on color, shape and size of the different parts of the tree.

- Seed storage and treatment.
- Seedling and nursery management activities.

- **Why and How buttons** - provide explanation on:
    - Why the question is asked.
    - Why the species are recommended.
    - What exactly the questions and choices mean.
    - How the answer is calculated.

- **Print button** -to print the current screen.
- **Help button** - to show the new user the functions of the system briefly.

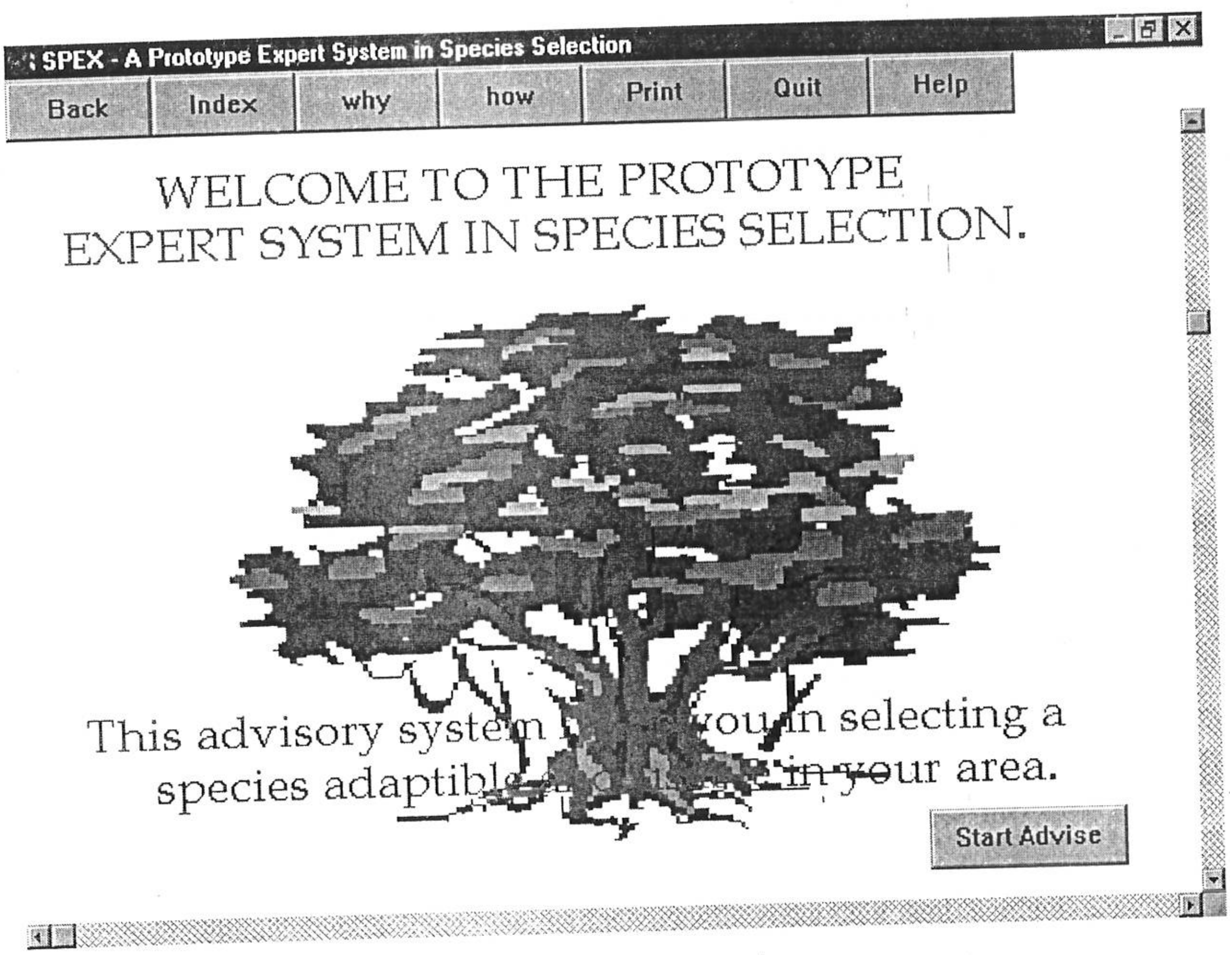


FIGURE 5.3 WELCOMING PAGE OF THE PROTOTYPE EXPERT SYSTEM, SPEX.

- With a click of the start advise button in the main page, the window shown below in figure 5.4 is opened. Two choices are presented in this window. The answer "yes" continues the question and answer session, but a "No" answer enables the user to exit the system.

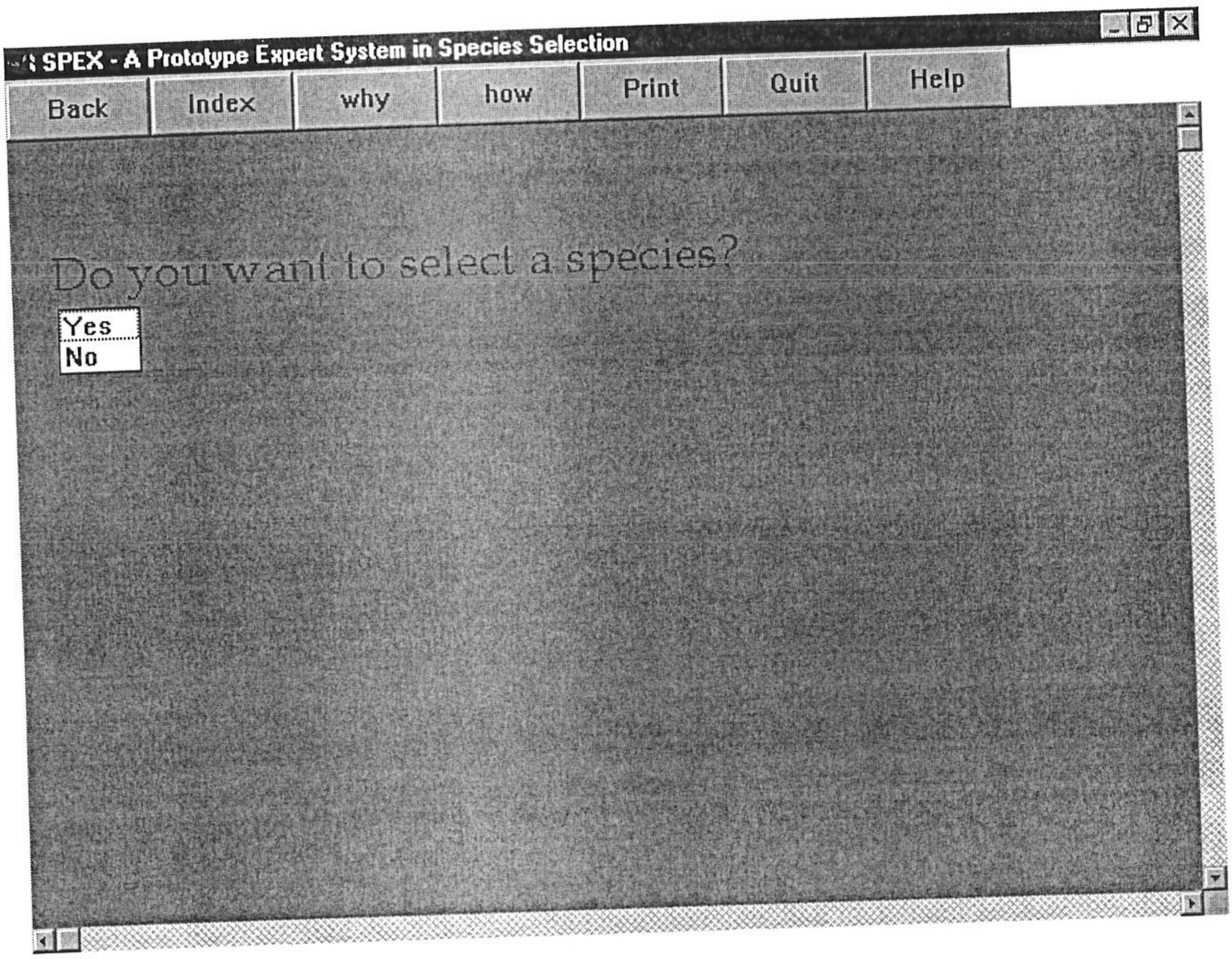


FIGURE 5.4 A CONFIRMATION WINDOW IN SPEX.

- A "yes" answer to the window in figure 5.4. Opens the next two windows, one after the other. Questions in these two windows pop up about the rainfall range and soil texture in the area/site, thereby inferring the site type.
- Then the system enquires about the land use system i.e., by giving two choices, cropland and non-cropland.

- In the subsequent windows, the system goes on to specifics until it gets the specific reason for the species to be selected. And the system narrows the species even further, as the user clicks options for the given questions. A sample question from the interface is presented in Figure 5.5.

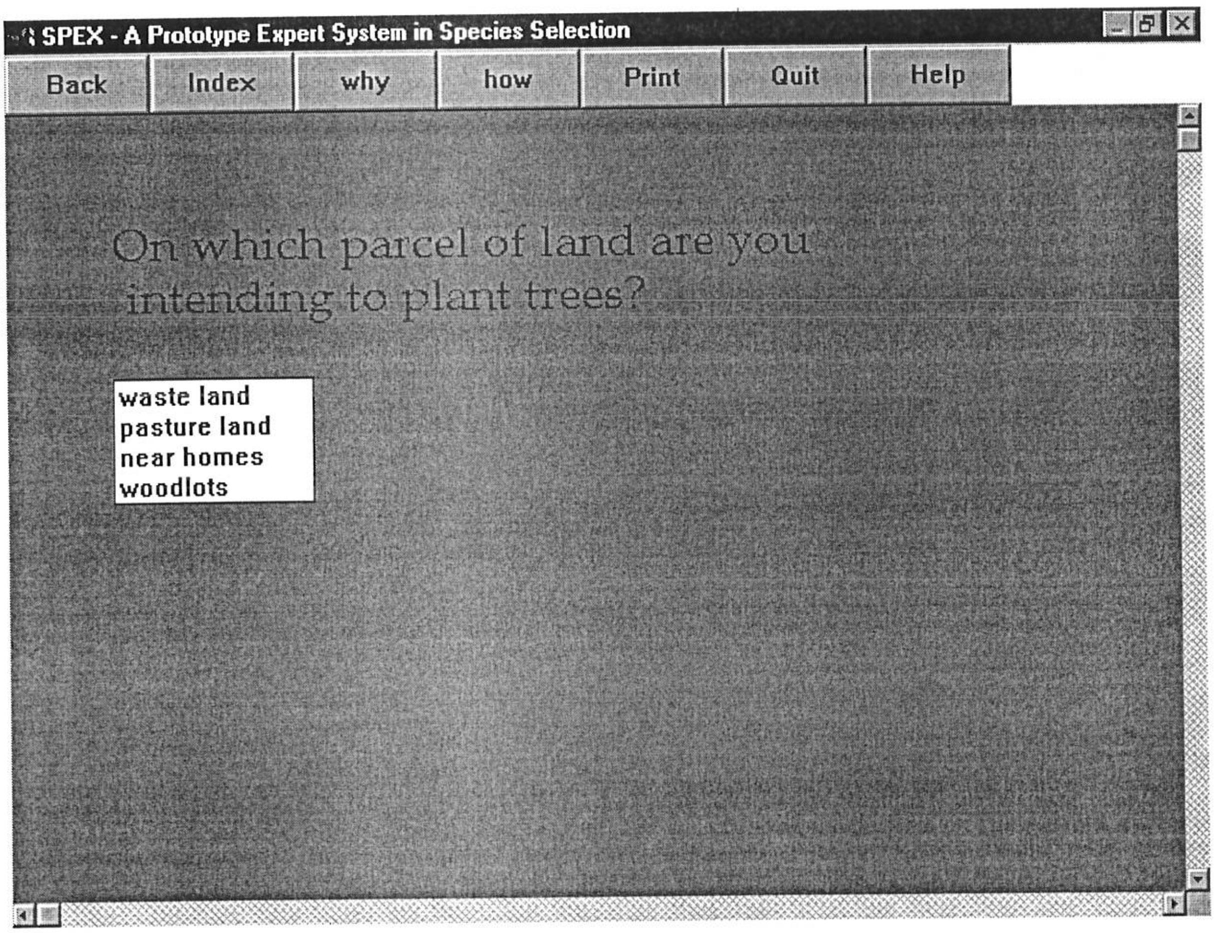


FIGURE 5.5 ONE OF THE QUESTIONS WITH CHOICES IN SPEX.

At last, the system outputs a recommendation (as shown below in figure 5.6), species with a combined Certainty Factors (CF), which shows the certainty of the species fitting in the required need of the user, given its appropriate site. The certainty factors are listed in a sorted manner to enable the user to consider species, which is higher in CF and appropriate for the site as compared to others. In the recommendation window, the button 'continue' leads the user to a confirmation window in figure 5.4.

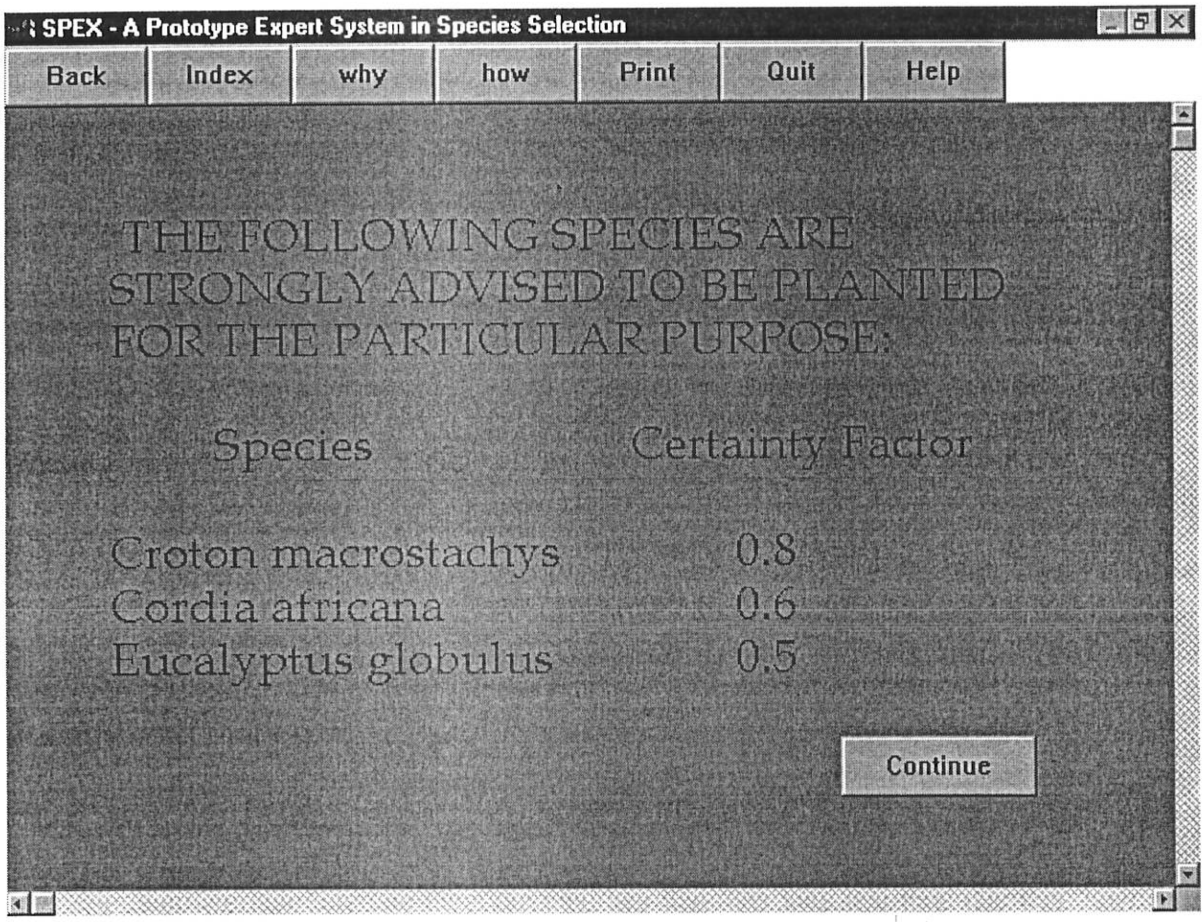


FIGURE 5.6 THE OUTPUT/ RECOMMENDATION WINDOW OF SPEX

5.7 TESTING AND EVALUATION

The Prototype Species Selection Expert System, SPEX, is checked first by the developer through reading and examining the production rules in the knowledge base. Thereafter, the prototype is debugged and confirmed that the system runs successfully.

When reading the rules the main issues taken into considerations are:

- Whether the rules do precisely/correctly what they have been intended to do without contradiction.
- Whether the rules simulate the expert's way of inferring during the communication with the user.

Rule contradictions are those statements in rules where conclusions are different when the premise (if clauses) is the same. To avoid such contradictions, the rules are carefully designed to have a hierarchy of statements as premises, as shown below, taking advantage of the hierarchical nature of the concepts.

```
if ?site is 'clay wet land'
 and ?'land use' is 'crop land'
 and ?'crop land' is soil
 and ? soil is 'soil fertility'
 and ?'soil fertility' is mulching
 then 'recommended species' is 'Albizzia lebbek (0.6) '.
```

So, in the 'if' statements shown above, rules are constructed or structured in the form of hierarchy. This means, the system asks the value of the rule statements hierarchically, one after the other, as indicated in the following example:

Value of land use is cropland
Value of cropland is soil
Value of soil is soil fertility
Value of soil fertility is mulching

Since the prototype system is an advisory system, two experts in the domain and four users (the Junior Researchers) of the system were asked to test whether the system simulates the session between the expert and users, and eventually gives correct advise that is meaningful and workable to the user. The general procedure followed in this respect is to ask the expert use the system and verify: -

- Whether the system solves the problem of species selection, in particular to the test cases given.
- Whether the knowledge base gives acceptable results.

For this purpose, experts and users formulate different test data based on the various user-expert sessions. The test data is prepared by letting the users/ experts choose the different options by themselves. The developer of the system then, screens out the user's request, by matching it with the rules and facts in the knowledge base before it is used as a test data. Besides, to make sure that the sites chosen are as different from the previously chosen, the users were told that a particular site has already been chosen, in times the site was chosen previously.

Among these, the following sample simulation is used in preparing test data submitted to the system (Other test data are attached in Appendix IV).

> *The user states that he needs a species adaptable to the area with rainfall of 1200 mm and, with loam textured soil. In addition, the user also asserts species to be planted in a non-cropland, specifically in woodlots to be used as energy source for the farmer in a form of fuel wood.*

Then, based on the user's need, experts and users prepare test data in the following way:

- Zone = semi-arid lowland, as rainfall is 1200mm (within range of 1050-1400 mm).
- Soil texture of the area= loam
- Land use = non cropland
- Problem with cropland components = need for energy source
- Species used in alleviating energy problem = firewood species.

When the expert/other users fed the above requirements to the system, SPEX outputs the following species on priority basis.

SPECIES	CERTAINTY FACTOR
Eucalyptus camaldulensis	0.8
Acacia saligna	0.7
Casuarina equistifolia	0.5
Azandrichta indica	0.4
Cupressus lustanica	0.4
Acacia nilotica	0.3

Table 5.3 Out put of SPEX for the test data.

In outputting the results, the Certainty Factor (CF) is arranged in a sorted manner. So, the user chooses the highest performing species according to the expert's belief of the species suitability in the specified area using the Certainty Factor. In this case, Eucalyptus camaldulensis is a better performing species than the other species listed above.

In times when the Certainty Factor of two or more species turn out to be the same, SPEX gives species descriptions to let the user decide and differentiate the species by their secondary uses which the user would like to plant. The following descriptions are taken based on table 5.3, where two species, Azandrachta indica and Casuarina equistifolia, have same CF (i.e 0.4).

Azandrachta indica - could also be used for contour strips, plantations along the roads, for fuel wood and charcoal, as fodder for bees (for honey & wax production), as shade and ornamental tree.

Casuarina equistifolia - can also be used as a conservation plant in terracing and in erosion control, plantations along the roads, and for firewood

SPEX displays only the benefits as shown above, but no Certainty Factor is attached for every benefit. This means that the user could not plant basing his decision only by the secondary benefits enumerated, but the above enumerations help him to distinguish between the species if he wants to plant one of them based on the primary benefit (i.e. windbreak). But as the prototype expert system developed further, it is advisable to include more choices of benefits/uses with some level of confidence, so that user could choose even from the secondary benefits.

Since the outset of the study, the objective of the experiment is not to come up with an operational and full-fledged expert system. Rather, it is an academic exercise done to demonstrate the potentiality of expert system to forestry sectors and to experiment in the tools and techniques of expert system. This has been shown by taking a task of species selection as applied to FRC. The results of the study showed that it is encouraging for further research to develop an applicable system. The statement is confirmed both by users and experts.

The experts commented that the system solves the problems in test cases successfully, after going through SPEX and its test results. The experts, then, conclude that the rules and the system as a whole are acceptable since they simulate the basic questions the expert asks before recommending a species to the site.

The users also confirmed the experts' comment that the system is workable and acceptable to the user, since the system narrows down users' demand, as experts do, the user could articulate his needs better. Additionally, the users asserted that the incorporation of Certainty Factor into SPEX is helpful, because users understand the degree of preference of experts for species, and the level of species performance in a site. The experts and users also suggested that the system should include more descriptions on species management. Finally, both users and experts recommended that the application of the prototype expert system could be expanded to other departments in the center through integration of expert system development within the computerization project held by FRC.

The system is evaluated by comparing the comments and test results found by experts (those closer with the system) and users (those with few knowledge of the system). It is observed that the comments and results between the users and experts are found to be matching greatly. The matching of comments of users and experts about the acceptability and applicability of the expert system shows that expert systems performance is acceptable in all sides, and it can be applied in species selection in particular and forestry in general.

5.8 DISCUSSIONS

It can be observed from the testing and the comments of the users and experts that the expert system is acceptable and applicable to species selection. This section discusses the major problems encountered in the course of the study, and the way they are solved.

- **Communication barrier** – the communication with experts in terms of procedures of research and species choice is minimal, because the developer is an outsider to forestry research; i.e. more explanation sessions were needed to elicit knowledge from experts. Secondly, the experts have hardly any knowledge in computers that it takes sessions to explain each expert how the computers, especially expert systems, could solve the problems they are facing. The developer, through reading books to get ideas and concepts involved in species selection and based on his educational background, has minimized such communication barriers.

- **Subjective assessment** – the procedures used by the expert is subjective, because of the observational nature of his research and the uncertainty involved in the experimental and statistical data available in the center. The subjectivity of the data has problems in that it involves bias. That means, the expert might assign a higher rank for those species he liked most, and an observation he had years ago might proof wrong as years went by, but he might stick to his belief that only his results are correct. Due to such problems, the researcher has insisted to give his reasoning for assigning a certainty factor to a species. As a result, the prototype incorporates the reasoning for assigning the certainty with more details about the species descriptions, as shown in the user interface.

- **Hierarchical knowledge structuring** – using hierarchy in knowledge structuring has a benefit of solving rule contradictions, flawed reasoning or questioning steps. But problems have been encountered in assigning the different properties of trees for every category of species usage in the hierarchy. As another option, these properties were put in the explanation facilities for the user to view the species properties involved in selecting a species for the right use.

During the research process it has been learned that the frame-based system could handle such system of categorizing the properties of a concept, even though knowledge pro could not support such knowledge representation scheme.

- **Combining CF** - Some of the combined CF in the rules were found to be very lower in this technique of calculating CF. This is because when combining the CF in the recommendation, multiplication of the site and usage CF are done. As a result, even though one of the CFs is very high, the lower CF of the other results in an even lower updated CF.

In Table 5.2, for instance, Juniperus procera species' best performance in site (0.9), decreases considerably, when it is multiplied with yield (0.6), with a resulting combined

CF of 0.5. The implication of the discovery in CF calculation is that the CF is a good way of limiting the primary list of species because it takes best performing species in their group. But when limited number of species is considered, as in the case of this study, their elimination from the list might affect the number of recommended species for the user to choose. Other techniques of handling uncertainty could also be considered and compared with the result found in this research.

5.9 DESIGN REQUIREMENTS

a) Hardware Requirement

Though, SPEX is developed in SISA laboratory using PENTIUM II PC with 32 MB system memory and 233 MHz processor speed, according to KpWin (1991), it is suggested that knowledge-pro for windows (KpWin) could be used in PCs having at least 4MB system memory and processor speed of 133 MHZ. However, 8MB memory at least for use with visual C++ environment is recommended.

b) Software Requirement

The knowledge pro for windows (KpWin) shell is version 3.0. To work on this version of KpWin, there is a need for at least the latest versions of Microsoft C/C++ (V7.0) or Borland C++ (V3.1), or the windows version of C++, i.e. Microsoft Visual C++.

In the network environment, the new system requires a network operating system, preferably Windows NT 4.0 and at least Windows 95 operating system on client PCs.

UNIT SIX
CONCLUSION AND RECOMMENDATIONS

6.1 CONCLUSION

Ethiopia has a variety of climatic and environmental conditions to sustain the woody plants (tree and shrubs). Although research on these plants is being done since the 60s, there have been considerable problems in disseminating the research output to users. (Husnia,1997). This brings about negative consequences of improper and non-optimal use of forest resources

The research and resulting transfer of knowledge about the planting site and specific benefits as well as other issues related with plantation management between researchers in the FRC and foresters, extension workers and farmers, is hampered mainly by the isolated research, poor information and communications technology (ICT) use, lack of enough experts in the field of species selection, and lack of research compilation, which are recognized as holdup for development of forestry sector in the national level(EARO,1997).

However, there has been recognition for automating the process of decision-making and knowledge transfer in FRC, using the latest IT components. So, the above problems and the recognition of IT by the organization led the researcher to the choice of expert systems in a specific task of species selection.

The present study is therefore made in an effort to design prototype expert systems in species selection, which assists the information provision mainly to agricultural extension workers and researchers.

The first step in developing SPEX, the knowledge acquisition process, is done using an interview as a main methodology with three of the FRC expert researchers in the field. A document review of handbooks and manuals used by the researchers has also been undertaken.

The concepts obtained during the knowledge acquisition were then modeled using hierarchical structuring technique. Based on the model, the knowledge is represented into a knowledge base by means of a rule-based system. The knowledge-pro expert system shell was used to set up a knowledge base and finally to run SPEX. The inference system of the knowledge-pro uses backward chaining, which is found to be appropriate with the rules and the ultimate goal, recommending species.

Through the user interface, an interaction with the system is accomplished by going through a well-defined Question and Answer (Q&A) session. In Q&A session, the user answers a series of questions by choosing among the options/choices provided by the system. Then, at the end of the session, the system recommends tree species to be planted by making use of a Certainty Factor (CF), which shows the belief the system has on his recommendations.

Lastly, in the course of the development of SPEX, the study has demonstrated the potentiality and applicability of expert systems in the field of forestry in general and in the task of species selection in particular.

6.2 RECOMMENDATIONS

In this study, an attempt is made to develop an expert system for species selection. The study indicates that the expert system is appropriate and essential in forestry. For further undertaking of research in the system, the following recommendations are given.

1. As SPEX is a prototype system, the subsequent development processes, i.e., further testing and validation/evaluation of the system is a necessity. For that matter, users, outside of the panel of experts who helped in the early development of the system should be intensively involved in the last stages of system development. This is to further guarantee that the system has acceptable advice and recommendations, which includes: -

   - Enlargement of the knowledge base, to enrich the knowledge.
   - Evaluation of explanation facilities of their completeness and clarity.
   - Fine-tuning of the expert system's role through the development stages by improving a checklist role to a full-fledged system that has same status of reasoning with an expert, and
   - Evaluating the need for integration with other forestry information systems.

2. The species selection expert system, SPEX, is a prototype system. So, only sample species were used to build SPEX. In order to complete the system, it is necessary to include the available and researched species where experts have knowledge about. But since, the number of tree species in Ethiopia is about 300, and to put all related information and knowledge, and to make the system complete. More resources and

time are required. Furthermore, attempts should also be made to develop similar expert systems in forestry domains, like forest management and site quality research.

3. The study used a Certainty Factor (CF) method because the research data is incomplete or unavailable. In tasks where data collection is easier, but prediction is still difficult, probabilistic methods may better be used after further research on the subject. In addition, as indicated in the discussion part, the combined/calculated CF is always lower than the site and usage CFs. Other techniques, such as the probability methods used to calculate the belief/certainty of an expert should also be assessed and used, to solve the ever increasing/ decreasing combinations of two or more uncertainties in the technique of certainty factor.

4. In naming the species the expert system used a scientific names, which are standard throughout the world. But due to a need in naming the species in different languages, the system described the species in Amharic and Oromo, because these languages are widely spoken in Ethiopia. So, the system should include as many naming of a tree species as possible in local languages, to give more familiarity of the species to the user during the interaction with the system. Other factors of plantation activities to complete the description should be added as well by further investigating especially the needs of the plantation project manager.

5. Due to time and resources limitations, the expert system is developed using the existing rules and facts. To make the system learn, and be dynamic, there should be a need to incorporate additional features of probabilistic methods. By learning, it means to be trained to learn rules- it does this by reacting to a set of inputs to produce a set of

easily predicted (or previously known) outputs. Probabilistic methods help in modifying the output, by combining the input probability information by users.

6. The prototype system used hierarchical and rule-based methods. As described elsewhere in the study, using these methods bring about ease in knowledge representation. And they are natural ways of expert reasoning. Given enough data for the task, more time for knowledge acquisition, an ambition to build complex expert system, and finance for purchasing software tools, the frame based and network representations can be used as substitutes for hierarchical method.

REFERENCE

AAAI. (2000) Expert **Systems**. at: www.aaai.org/pathfinder/html .

Alison (1994) **Expert Systems**. at: www.cee.hw.ac.uk/~alison/ai3notes/chapter2_5.html.

Berhanu H., Nyalun, J. and Mustanoja, K (1988) **Forestry Research: Problems and Potentials Programme for Future Action**, Addis Ababa: Natural Resources Conservation and Development, Ministry of Agriculture.

Berhanu M. (1997) **Critical Aspects of State Regional Forest Resources Management in Ethiopia**, Ethiopian Forest Policy Workshop, and Addis Ababa: Ministry of Agriculture.

Benchimol, L, and Pomerol G, Levine P and Pomerol J. (1987) **Developing Expert Systems For Business**, Wiltshire: Anthony Rowe Ltd.

Boland, D. (1985) **Selection of Species and Provenances for Tree Introduction**, IDRC. at: (www.idrc.ca/library/document/074940)

Brule, J. (1986) **Artificial Intelligence: Theory, Logic and Application**, Pennsylvania: TAB Books

Cercone, N and Mccala, G. (1984) **AI: Underlying Assumptions and Basic Objectives,** Journal of The American Society of Information Science (JASIS), 35(5): 280-290.

CLAES (1996) **Expert Systems in Agriculture**. at: Http://Poteto.Cleas.Sci.Eg/Cleas/Citex/Citex.Html

Combs, M. (1984) **Developments in Expert Systems**: A Special Issue. The International Journal of Man-Machine Studies, London: Academic Press.

Cremer, K. (1990) **Trees for Rural Australia**, Melbourne: Inkata Press.

Drenth, H., Morris A, and Gweyneth T. (1991). **Expert Systems as Information Intermediaries**, Annual Review of Information Science and Technology (ARIST), Volume, 26, New Jersey: Learned Information Inc.

EARO (1997) **Forestry Strategic Plan**, Addis Ababa: Ethiopian Agricultural Research Organization (EARO).

Ebrahim K. (1999) **An Expert Based Admission and Placement in Addis Ababa University,** Addis Ababa: school of information studies for Africa, Addis Ababa university.

EFAP (1994) **The Challenge For Development**, Volume II, and Final Report, Addis Ababa: Ethiopian Forestry Action Program (EFAP).

EPA (1997) **Environmental Policy**, Addis Ababa: Environmental Protection Authority.

Englemore, S. (Ed) (1993) **Knowledge- Based Systems in Japan,** Japanese Technology Evaluation Center at: http://itiri.loyola.edu/kb/

Feigenbaum, E and McCorduck, P. (1983) **The 5th Generation: AI and Japan's Computer Challenge to the World,** Massachusetts: Addison- Wesley Publishing.

Feltovich, P, Ford K, and Hoffman R. (1997) **Expertise in Context**, California: American Association for Artificial Intelligence.

FRC (1986) **Growth of Some Forest Trees in Ethiopia and Suggestions for Species Selection in Different Climatic Zones**, Addis Ababa: Forestry Research Center.

Amare Getahun, Abate, Feyisa And Gebre, Belay (1990) **A Brief History Of FRC and Forestry Research in Ethiopia - 1890-1990,** Addis Ababa: Forestry Research Center.

Harmon, P. and King, D. (1985) **Expert Systems: Artificial Intelligence in Business,** New York: Wiley.

Husnia I. (1997). **Australian Species in African Plantations Country Reports: Ethiopian Case**, Ottawa: IDRC. at:- www.idrc.ca/library/document/074949/chap23_e.html)

Inneson, P.(1985) **Expert Systems: an Application in Soil Research ,in Computers and Forestry,** New York: John Wiley and Sons.

IPT (2000) **Artificial Intelligence: An Overview.** at: www.sunnybank.qld.edu.au/ipt/ai/ai3.htm

Kandel, A (1991). **Fuzzy Expert Systems,** London: CRC Press.

KpWin (1991) Knowledge Pro for Windows: A User Manual, New York: Knowledge Garden, Inc.

Liebowitz, J. (1988) **Introduction to Expert Systems**, Santa Cruz, CA: Mitchell Publishing, Inc.

Masuch, M. (Ed.) (1990) **Organization, Management, and Expert Systems**, Berlin: Walter De Gruiter.

MOA (1992) **Workshop on Exotic Species of Ethiopia: WURC and Munessa Shashemenne State Forest Development Project**, Addis Ababa: Ministry of Agriculture.

MOA (1999) **National Agricultural Information Systems (NAIS) Needs Assessment and Systems Design**, Draft Document, Addis Ababa: Ministry of Agriculture.

MOA (2001) **Forest Resources Assessment (FRA) 2000 for Ethiopia**, Addis Ababa: Ministry of Agriculture.

Pancel, L. (1992) **Handbook for Forestry Practices**, Ottawa: IDRC.

Parsaye, K and Chignell, M. (1988) **Expert Systems for Experts,** New York: John Wiley & Sons.

Pederson, K. (1989) **Expert Systems Programming: Practical Techniques for Rule Based Systems,** New York: John Wiley & Sons.

Richardson, J. (1995) **Knowledge-Based Systems for General Reference Work: Applications, Problems, and Progress,** San Diego: Academic Press.

Rocheleau D, Weber F, and Field-Juma A. (1988) **Agroforestry in Dry Land Africa,** Nairobi: ICRAF.

Schmoldt, D. (1986) **Expert Systems in Forestry: Utilizing Information and Expertise for Decision Making,** Amsterdam: Elsevier Science Publishers.

Schmoldt D. and Martin G. (1989) **Development and Evaluation of an Expert System for Diagnosing Pest Damage of Red Pine in Wisconsin,** Forest Science, 35(2), Pp. 364-387.

Stefik, M. (1998) **Introduction to Knowledge Systems,** California: Morgan Kaufmann Publishers.

Wormell, I. (Ed.) (1987) **Knowledge Engineering: Expert Systems and Information Retrieval.** London: Taylor Graham.

WURC (1995) **Commercial Timbers of Ethiopia: Research Report,** Technical Bulletin No 2, Revised and Enlarged Edition, Addis Ababa: Wood Utilization Research Center.

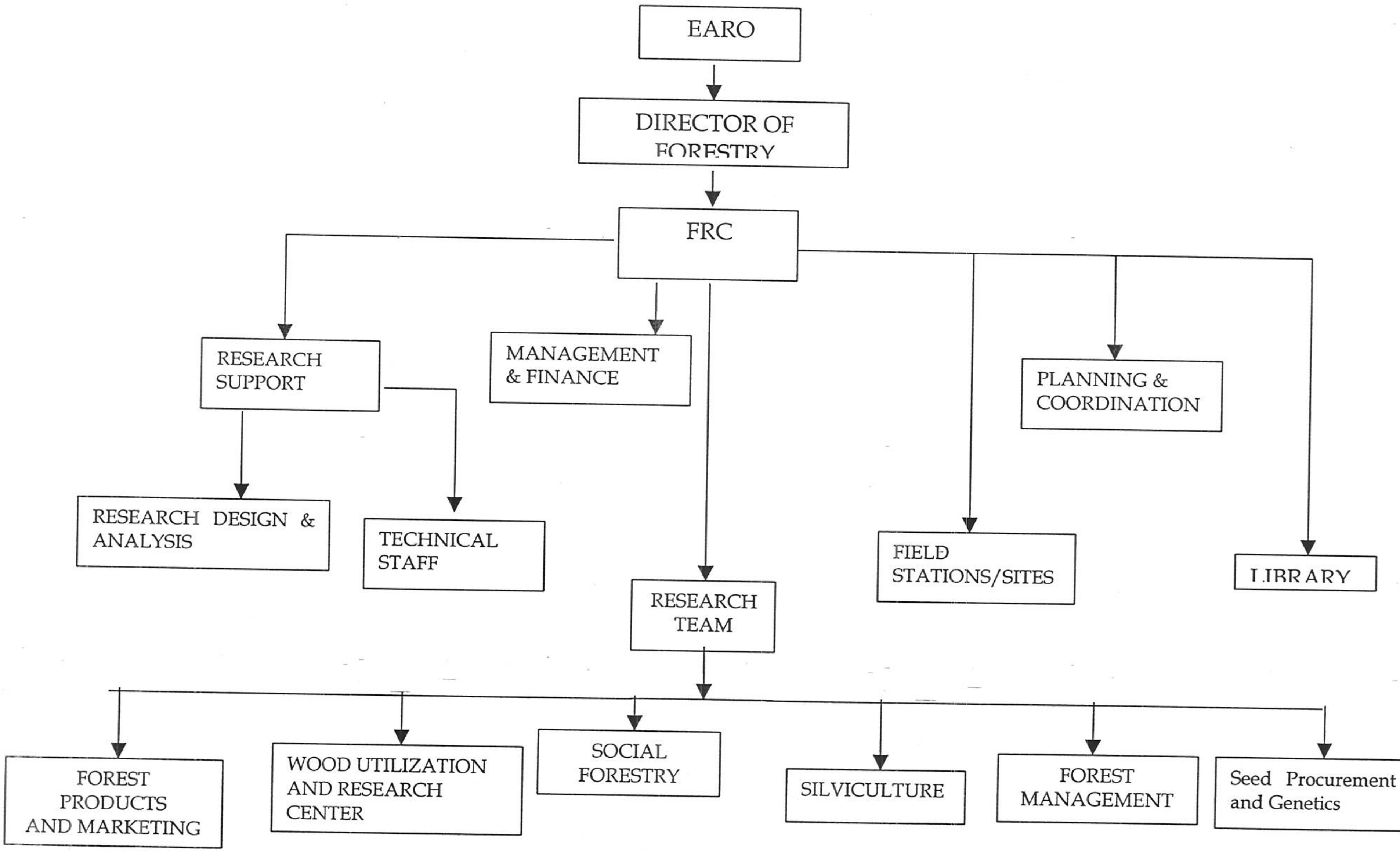
Yaghamai, S, and Maxin, J. (1984) **Expert Systems: A Tutorial,** Journal of The American Society of Information Science, 35(5), New York: John Wiley & Sons.

Zahedi, F. (1993) **Intelligent Systems For Business: Expert Systems with Neural Networks,** California: Wadsworth Publishing Co.

APPENDICES

APPENDIX I.

THE ORGANIZATIONAL STRUCTURE OF FRC



APPENDIX II.

SAMPLE RULES USED FOR THE SPECIFIC SITE

SITE RULES

****

```
topic  site.
  set_number_of_values (site,1).

if ?rainfall is '700-1050'
and ?'soil texture' is loam
then site is 'loam semi-arid lowland'
```

****

END-USE RULES

****

```
topic  'recommended species'.
set_number_of_values ('recommended species',1).

if  ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is soil
and  ?soil is 'soil fertility'
and  ?'soil fertility'  is 'nitrogen fixation'
then 'recommended species' is  '#n#t Juniperus procera#t(0.5) #n#tEucalyptus saligna#t(0.5) #n#tHagenia abyssinica#t(0.4)#n#tEucalyptus globulus#t(0.4) #n#t Gravellea robusta#t(0.3) '.


if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is soil
and  ? soil is 'soil fertility'
and ?'soil fertility' is mulching
then 'recommended species' is '#n#t Gravellea robusta#t(0.6)##n#t Juniperus procera#t(0.5) #n#t Cordia africana#t(0.5) #n#tHagenia  abyssinica#t(0.4) #n#tEucalyptus saligna#t(0.4) '.

if  ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is soil
and  ?soil  is 'soil retension'
then  'recommended  species'  is  '#n#t  Juniperus  procera#t(0.5)#n#t  Gravellea robusta#t(0.5)#n#tHagenia  abyssinica#t(0.5)  #n#t  Cordia  africana#t(0.4)  #n#tCroton macrostachys#t(0.3)'.


if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is 'agricultural crop'
and  ?'agricultural crop' is 'windbreak'
```

then 'recommended species' is '#n#t Cordia africana#t(0.7) #n#tPinus patula#t(0.3)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is 'agricultural crop'
and ?'agricultural crop' is 'shade and shelter'
then 'recommended species' is '#n#t Gravellea robusta#t(0.7)#n#tHagenia abyssinica#t(0.6) #n#t Cordia africana#t(0.6) #n#tCroton macrostachys#t(0.5) #n#tPinus patula#t(0.5) #n#tEucalyptus globulus#t(0.21) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is 'agricultural crop'
and ?'agricultural crop' is 'yield increment'
then 'recommended species' is '#n#tCroton macrostachys#t(0.7) #n#t Juniperus procera#t(0.4) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is 'sloped landscape'
and ?'sloped landscape' is 'terracing'
then 'recommended species' is '#n#t n#t Gravellea robusta#t(0.6) #n#tEucalyptus saligna#t(0.4) #n#tCroton macrostachys#t(0.3) #n#tPinus patula#t(0.3)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is 'sloped landscape'
and ?'sloped landscape' is 'contour strips'
then 'recommended species' is '#n#tCroton macrostachys#t(0.6) #n#t Cordia africana#t(0.5) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is 'sloped landscape'
and ?'sloped landscape' is 'small structures'
then 'recommended species' is '#n#tCroton macrostachys#t(0.3)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'waste land'
and ?'waste land' is 'gully reclaimation'
then 'recommended species' is '#n#tCroton macrostachys#t(0.7) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'waste land'
and ?'waste land' is 'swamp reclaimation'
then 'recommended species' is '#n#tEucalyptus saligna#t(0.6) #n#t Eucalyptus globulus#t(0.5) '.

then 'recommended species' is '#n#t Cordia africana#t(0.7) #n#tPinus patula#t(0.3)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is 'agricultural crop'
and ?'agricultural crop' is 'shade and shelter'
then 'recommended species' is '#n#t Gravellea robusta#t(0.7)#n#tHagenia abyssinica#t(0.6) #n#t Cordia africana#t(0.6) #n#tCroton macrostachys#t(0.5) #n#tPinus patula#t(0.5) #n#tEucalyptus globulus#t(0.21) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is 'agricultural crop'
and ?'agricultural crop' is 'yield increment'
then 'recommended species' is '#n#tCroton macrostachys#t(0.7) #n#t Juniperus procera#t(0.4) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is 'sloped landscape'
and ?'sloped landscape' is 'terracing'
then 'recommended species' is '#n#t n#t Gravellea robusta#t(0.6) #n#tEucalyptus saligna#t(0.4) #n#tCroton macrostachys#t(0.3) #n#tPinus patula#t(0.3)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is 'sloped landscape'
and ?'sloped landscape' is 'contour strips'
then 'recommended species' is '#n#tCroton macrostachys#t(0.6) #n#t Cordia africana#t(0.5) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'crop land'
and ?'crop land' is 'sloped landscape'
and ?'sloped landscape' is 'small structures'
then 'recommended species' is '#n#tCroton macrostachys#t(0.3)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'waste land'
and ?'waste land' is 'gully reclaimation'
then 'recommended species' is '#n#tCroton macrostachys#t(0.7) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'waste land'
and ?'waste land' is 'swamp reclaimation'
then 'recommended species' is '#n#tEucalyptus saligna#t(0.6) #n#t Eucalyptus globulus#t(0.5) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'pasture land'
and ?'pasture land' is 'fodder for animals'
then 'recommended species' is '#n#t Cordia africana#t(0.4) #n#tCroton macrostachys#t(0.4) #n#t Gravellea robusta#t(0.3) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'pasture land'
and ?'pasture land' is 'resistant species'
and ?'resistant species' is draught
then 'recommended species' is '#n#tPinus patula#t(0.4)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'pasture land'
and ?'pasture land' is 'resistant species'
and ?'resistant species' is 'pest resistant'
then 'recommended species' is '#n#t Juniperus procera#t(0.7)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'near homes'
and ?'near homes' is 'in-between planting'
and ?'in-between planting' is 'living fences'
then 'recommended species' is '#n#tCroton macrostachys#t(0.6)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'near homes'
and ?'near homes' is 'in-between planting'
and ? 'in-between planting' is 'roadside planting'
then 'recommended species' is '#n#t Juniperus procera#t(0.3)#n#tPinus patula#t(0.3) #n#tEucalyptus globulus#t(0.21) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'near homes'
and ?'near homes' is 'income generating'
and ?'income generating' is 'honey & wax production'
then 'recommended species' is '#n#tHagenia abyssinica#t(0.7) '.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'woodlots'
and ?'woodlots' is energy
and ?energy is 'fuel wood'
then 'recommended species' is '#n#t Gravellea robusta#t(0.7)#n#tEucalyptus globulus#t(0.6) #n#tPinus patula#t(0.5) #n#t Juniperus procera#t(0.3)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'woodlots'
and ?'woodlots' is energy
and ?energy is charcoal
then 'recommended species' is '#n#tEucalyptus globulus#t(0.6) #n#tPinus patula#t(0.3) #n#t Juniperus procera#t(0.2)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'woodlots'
and ?'woodlots' is timber
and ?timber is 'agricultural tools'
then 'recommended species' is '#n#tHagenia abyssinica#t(0.7) #n#tEucalyptus saligna#t(0.7) #n#tEucalyptus globulus#t(0.7) #n#tPinus patula#t(0.4) #n#t Cordia africana#t(0.4) #n#tCroton macrostachys#t(0.4)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'woodlots'
and ?'woodlots' is timber
and ?timber is 'poles & posts'
then 'recommended species' is '#n#tCroton macrostachys#t (0.5) #n#t Cordia africana#t(0.4) #n#tEucalyptus globulus#t(0.3) #n#tHagenia abyssinica#t(0.2) #n#tPinus patula#t(0.24)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'woodlots'
and ?'woodlots' is timber
and ?timber is 'construction materials'
then 'recommended species' is ' #n#tEucalyptus saligna#t(0.6)#n#tEucalyptus globulus#t(0.5) #n#tCroton macrostachys#t(0.5) #n#tHagenia abyssinica#t(0.4) #n#t Cordia africana#t(0.3) #n#tPinus patula#t(0.3)'.

if ?site is 'loam semi-wetland'
and ?'land use' is 'non crop land'
and ?'non crop land' is 'woodlots'
and ?'woodlots' is timber
and ?timber is furniture
then 'recommended species' is '#n#tHagenia abyssinica#t(0.6) #n#tEucalyptus saligna#t(0.6)#n#tEucalyptus globulus#t(0.6) #n#tPinus patula#t(0.4)'#n#t Cordia africana#t(0.3)'.

APPENDIX III

QUESTIONS USED IN STRUCTURED INTERVIEWS

Questions used in interviews, which are dealt in continuous sessions with each expert are: -

- What are the objectives of the section with regard to the site selection task?
- What are the concepts/factors involved in site selection of species?
- What are the symbolic /numeric values for the concepts involved here?
- How many of the theoretical factors do you use in a process of reaching recommendations? Why?
- Could you use these sample species only, and differentiate between them by ranking according to their performance in (*name of factors*), and using these concepts?
- Explain how a species are ranked over the others?
- Could you assign value for the ranking (to prioritize among species)? Why?
- How do you differentiate species having similar value ranking?

APPENDIX IV

THE TEST DATA

The test data prepared during testing, along with the output of SPEX, when it is fed to the data, are presented below.

1. Zone = semi-arid lowland ,

   Because rainfall chosen was in range 700-1050mm

   Soil type = sandy

   land use = non crop land

   problem = need for resistant species in pasture land

   specific problem a species resists in the area = draught resistance

when the expert inputs the above requirements, the system recommends the following species:

Species	Certainty factor
Acacia tortillas	0.4
Casuarina equistifolia	0.4

2. Zone = loam semi-wetland

   Because rainfall chosen was in range 1050-1400mm

   Soil type = loam

   land use = crop land

   problem = need to increase the fertility of the soil in the farm.

   specific way of increasing fertility to the soil = adding leaf mulch by the species.

when the expert inputs the above requirements, the system recommends the following species on priority basis:

Species	Certainty factor
Juniperus procera	0.6
Gravellia robusta	0.6
Cordia africana	0.5
Hagenia abyssinica	0.4
Croton macrostachys	0.4

3. Zone = wet region

Because rainfall chosen was in range 1400mm

Soil type = clayey

land use = crop land

problem = a species used for terracing practice in sloped landscape.

when the expert inputs the above requirements, the system recommends the following species on priority basis:

Species	Certainty factor
Cupressus lustanica	0.6
Gravellia robusta	0.5
Croton macrostachys	0.4
Juniperus procera	0.4
Eucalyptus camaldulensis	0.4

4. Zone =loam dry lowlands

Because rainfall chosen was in range 350- 700mm

Soil type = loam

land use = non crop land

Problem = need for energy source.

Specific way of solving energy shortage =fuel wood (firewood) species.

when the expert inputs the above requirements, the system recommends the following species on priority basis:

Species	Certainty factor
Eucalyptus camaldulensis	0.8
Acacia saligna	0.6
Cupressus lustanica	0.4
Azandrichta indica	0.4

5. Zone = dry lowland

Because rainfall chosen was in range 350-700mm

Soil type = sandy

land use = crop land

Problem = need to protect crops from light damage.

Specific way of protecting from light =choosing shade and shelter species.

when the expert inputs the above requirements, the system recommends the following species on priority basis:

Species	Certainty factor
Azandrichta indica	0.7
Acacia albida	0.4
Albizzia lebbek	0.3

6. Zone = dry lowland

Because rainfall chosen was in range 350-700mm

Soil type = clay

land use = non crop land

Problem = need for timber

Specific timber use sought = timber for making up furniture.

when the expert inputs the above requirements, the system recommends the following species on priority basis:

Species	Certainty factor
Hagenia abyssinica	0.5
Acacia albida	0.3

Declaration

This thesis is my original work and has not been submitted as a partial requirement for a degree in any other university.

Samir Abduselam Ibrahim
July 2001

The thesis has been submitted for examination with our approval as university advisors.

Tesfaye Birru
July 2001

Million Meshesha
July 2001