

USING AN OFF-THE-SHELF SHELL TO DEVELOP AN INCOME TAX
EXPERT SYSTEM

Dirk Udo Els

A Research Report Submitted to the Faculty of Commerce,
University of the Witwatersrand, Johannesburg, in
Partial Fulfillment of the Requirements for the Degree
of Master of Commerce (by Coursework).

Johannesburg, December 1988

DECLARATION

I declare that this research report is my own, unaided work. It is being submitted in partial fulfillment of the requirements for the degree of Master of Commerce (by Coursework). It has not been submitted before for any degree or examination in any other University.

Dirk Udo Els

Dirk Udo Els

9 day of July, 1988

ABSTRACT

Expert systems are computer systems which simulate problem-solving behaviour of human experts, thus reaching similar conclusions. Commercial applications of expert systems are not yet widespread. This research paper documents by way of a case study, the development of FOREX, an expert system in the domain of income tax, specifically the tax implications of transactions in foreign currencies. The objective of the research is to provide practical guidelines in the areas of knowledge acquisition and elicitation, the encoding and structuring of the elicited knowledge, and the utility of using expert system shells. The paper reviews the concepts of expert systems and existing income tax expert systems are examined. The domain of FOREX is discussed. The knowledge acquisition and elicitation process used in the development of FOREX is described as well as the methods used for encoding the knowledge. The utility of using off-the-shelf shells is discussed. The methods used to validate the system are also described. This is followed by a sample consultation with the completed system. The paper concludes with the findings of the research relating to tax-based expert systems, the domain of income tax, knowledge acquisition and elicitation, the encoding of knowledge, the utility of using shells and the validating of systems.

CONTENTS

	Page
Abstract	
1.Introduction	1
2.Research Objective and Methodology	4
3.Expert Systems	7
3.1.Definition	7
3.2.Components	8
3.2.1.The knowledge base	
3.2.2.The inference engine	
3.2.3.The explanatory interface	
3.2.4.The knowledge-acquisition module	
3.3.Suitable tasks for expert systems	11
3.4.Advantages and benefits	13
3.5.Disadvantages and limitations	14
4.Existing Income Tax Expert Systems	16
5.The Knowledge Domain: South African Income Tax	22
5.1. Identifying sources of knowledge	22
5.2. Basis of taxation	23
5.3. Applicable sections of the Act	24
5.4. Other knowledge sources	26
5.5. Nature and types of problems	28
6.The Knowledge Acquisition Process	30
6.1. Literature survey	30
6.2. Identifying the experts	34
6.3. Extracting the knowledge	36
6.4. Analysis of knowledge elicited	37
6.5. Knowledge validation	39
6.6. Prototype system	40
6.7. Prototype refinement	40

	Page
6.8.Findings	41
6.8.1.Identification of the experts	
6.8.2.Knowledge elicitation	
6.8.3.Countering bias	
6.8.4.Domain expertise of knowledge engineer	
6.8.5.Further research	
7.Encoding the Knowledge	45
8.Utility of Using an Off-The-Shelf Shell	50
8.1.Shell selection	50
8.2.VP-Expert	50
8.3.Experience with VP-Expert	51
8.4.Advantages of using a shell	52
9.Validating the System	54
9.1.Approaches to validation	54
9.2.Validating FOREX	58
10.Sample Consultation	61
11.Concluding Findings	65
11.1.Tax based expert systems	65
11.1.1.The domain of income tax	66
11.3.Knowledge acquisition and elicitation	66
11.3.1.Expert identification	
11.3.2.Knowledge elicitation techniques	
11.3.3.Validation of elicited knowledge	
11.3.4.Practitioners as system developers	
11.4.Encoding of acquired knowledge	68
11.5.Utility of using a shell	68
11.6.Validating the system	68
11.7.Summary	69
11.8.Further possibilities	69
References	61
Appendix A	

1. INTRODUCTION.

The concept of expert systems is becoming an increasingly more popular topic in South Africa among people who are involved primarily in the information technology field. For example, a major conference on expert systems was recently held at the Council for Scientific and Industrial Research in Pretoria (Computer Mail, 1988). Also the computer society of South Africa has created a special interest group in Artificial Intelligence which has been organising regular, open meetings since the beginning of 1986.

Expert systems are computer systems which attempt to simulate the problem-solving behaviour of humans and so reach similar conclusions as would human experts within a particular field of knowledge. The concept of expert systems arose out of research into Artificial Intelligence (AI) during the 1970's (Forsyth, 1986). Winston (op cit Lucchi et al., 1986) defines AI as "the study of ideas which enable computers to do the things that make people seem intelligent". Other technologies flowing from AI are, for example, robotics and speech generation and recognition.

A number of expert systems developed have proven to be highly successful. Certain examples are quoted time and

again in the literature. These usually include the systems RI/XCON and Dipmeter Advisor, details of which are discussed by Bobrow et al. (1986), MYCIN and DENDRALL (refer Pollitzer and Jenkins (1985) for brief descriptions) and XSEL (McDermott, 1982). While systems such as the above, are always mentioned in the literature, some authors such as Martins (1984) argue that expert systems are oversold. Martins says for example, that for applications of modest complexity, most expert system code is hard to understand, to debug, extend and maintain. He also states that current off-the-shelf tools (shells) cost too much, that they are poorly supported, that they lack adequate documentation, are hard to use and that they have limited applicability to complex problems.

This paper documents the development of an expert system in income tax using an off-the-shelf shell. The development of the system is discussed from the point of view of a non-expert in the field of expert system technology. The research report has three emphasis:

1. It highlights the practicality of using expert system shells.
2. It discusses the practical problems encountered in the knowledge acquisition process.
3. It discusses the practical problems encountered in the knowledge encoding process.

The research report first discusses the research objec-

tive and the methodologies utilized in carrying out the research. The paper then discusses the concept of expert systems. What an expert system is, the components of an system, its advantages and benefits, and the disadvantages and limitations of such systems are discussed. Following this, current systems in the domain of income tax which have been identified from a survey of the current literature, are examined. The paper then documents the knowledge domain of the system developed, namely South African income tax with specific reference to transactions in foreign currencies. The knowledge acquisition process utilized in the development of the system is then discussed. This is followed by a description of the methods used to encode the knowledge and a discussion of the problems encountered in using an off-the-shelf shell for the system. The methods used to validate the system are then described. The paper ends with a sample consultation using the system developed and concludes with a summary of the problems encountered and the methods developed to overcome the problems.

2. RESEARCH OBJECTIVE AND METHODOLOGY

To date most expert systems have been developed by researchers and academics whose area of expertise is in the field of information technology. In order for the concept of expert systems to realize its full potential, it is necessary to obtain the support of people outside the field of information technology. These will be the persons who will make use of such systems in their daily tasks. To make available the technology to such people, the practical difficulties of constructing and using expert systems must be identified and solved. This would then enable persons not au fait with the theory of expert systems to employ the technology for their own use.

The purpose of this research is to provide practical and general guidelines to others who wish to develop expert systems for their own use in any area of expertise. The objective of the research therefore is to identify problems which can arise when developing a practical expert system application and to provide solutions to such problems where possible.

The focus of the research was successively on:

1. The elicitation of the knowledge in the domain of income tax.

7. RESEARCH OBJECTIVE AND METHODOLOGY

To date most expert systems have been developed by researchers and academics whose area of expertise is in the field of information technology. In order for the concept of expert systems to realize its full potential, it is necessary to obtain the support of people outside the field of information technology. These will be the persons who will make use of such systems in their daily tasks. To make available the technology to such people, the practical difficulties of constructing and using expert systems must be identified and solved. This would then enable persons not au fait with the theory of expert systems to employ the technology for their own use.

The purpose of this research is to provide practical and general guidelines to others who wish to develop expert systems for their own use in any area of expertise. The objective of the research therefore is to identify problems which can arise when developing a practical expert system application and to provide solutions to such problems where possible.

The focus of the research is successively on:

1. The elicitation of the knowledge in the domain of income tax.

- 2.The representation of the elicited knowledge.
- 3.The utility of using a: off-the-shelf expert system shell the construction of the system.

The research thus addresses the following three broadly defined areas:

1.The knowledge acquisition process:

Which methods and techniques can be used to acquire the necessary knowledge? What problems arise when acquiring the expert's knowledge using the methods and techniques selected?

2.Structuring the acquired knowledge:

What methods should be used for structuring the acquired knowledge for the purpose of encoding the expert system knowledge base? What problems arise when structuring the elicited knowledge?

3.Using expert system shells:

What problems are encountered when using an off-the-shelf shell? What are the limitations of such a shell?

The research therefore provides an interface between the theory and practical applications. This report does not address the problems of hardware and software selection for expert system applications.

The research was undertaken by way of a case study. The case study documents the development of an expert system within a knowledge domain in which the author has some

experience, namely South African income tax. The expert system was developed within a specific sub-domain namely the income tax implications of transactions in foreign currencies. The system, named FOREX, was developed by the author using the expertise of chartered accountants employed by the Department of Finance, Inland Revenue.

The research report documents the methods used to acquire the necessary knowledge for the system, how the knowledge was encoded, what problems and limitations arose when using an expert system shell and how the system was validated. A sample consultation is also documented.

It is hoped that this research will contribute to the development of commercially viable expert systems. Not much has been written on the South African experience in this regard. This research report is therefore a contribution, specifically with regard to the practical problems of developing working and viable expert systems.

Finally it is hoped that this research may, in its own small way, contribute to alleviating the problem of the shortage of skilled manpower in South Africa, by encouraging the development of expert systems and so making available experts' knowledge to a much wider audience.

3. EXPERT SYSTEMS

As stated previously, expert systems are a resultant technology of initial research into the concept of Artificial Intelligence. What exactly are expert systems, what do they do, how are they constructed, what are their advantages, their disadvantages and limitations? This chapter briefly discusses the concepts of expert systems. The definitions and features of expert systems are examined. This is followed by a discussion of the structure and components of a typical expert system. Thereafter the tasks which expert systems perform are looked at. This is followed by a discussion of the advantages of expert systems over conventional programs and the benefits such systems hold in. Finally the disadvantages and limitations are examined.

3.1. Definition

Buchanan and Duda (1983) define an expert system as being "a computer program that provides expert-level solutions to important problems and that is:

1. Heuristic: - Reasons with judgemental and formal knowledge.
2. Transparent: - Provides an explanation of its line of reasoning.
3. Flexible: - Integrates new knowledge incrementally."

Adding to this definition, Forsyth (1985) has identified the following checklist of features that expert systems may have:

1. Expert systems are limited to relatively narrow domains of expertise.
2. They reason with uncertain data and unreliable rules.
3. They explain their way of reasoning in a comprehensible way.
4. The facts and inference mechanism are separate.
5. The systems grow incrementally.
6. The systems are usually rule-based.
7. They deliver advice as their output.

3.2. Components

A expert system can be seen to be a computer program which consists of a number of semi-autonomous components. Forsyth (1985) gives a effective summary. [Refer also to Hayes-Roth (1985) and Luconi et al. (1986) for graphic descriptions.] An expert system consists of:

1. A knowledge base.
2. An inference engine.
3. An explanatory interface.
4. A knowledge acquisition module (optional).

3.2.1. The knowledge base

This is one of the two fundamental components of an expert system. The knowledge base contains the essential information about the domain of expertise. This usually

consists of the expert's knowledge, judgemental and experiential, which is encoded in appropriate forms. The knowledge usually comprises the following elements:

1. Domain terms: - The technical words and jargon used by the experts in the domain.
2. Structural relationships: - The interactions of component entities within the domain.
3. Causal relationships: - The cause-effect relationships between components in the domain.

Various methods are used to encode this type of information. There are four main methods:

1. If-Then rules: - The knowledge is encoded in the form of condition/conclusion structures. For example, a rule encoding medical information might look like this;

```
IF temperature = 38 and patient = vomiting
THEN patient = sick.
```

2. Semantic nets: - The relations among objects are represented by links between nodes: e.g. cats are mammals.
3. Frames: - The knowledge about an object is presented in the form of a generalised record structure (i.e. frame format).
4. Horn clauses: - This is a form of predicate logic on which PROLOG (a programming language) is based.

Most systems developed to date make use of the If/Then rule format. Hayes-Roth (1985) says that this is

consists of the expert's knowledge, judgemental and experiential, which is encoded in appropriate forms. The knowledge usually comprises the following elements:

1. Domain terms: - The technical words and jargon used by the experts in the domain.
2. Structural relationships: - The interactions of component entities within the domain.
3. Causal relationships: - The cause-effect relationships between components in the domain.

Various methods are used to encode this type of information. There are four main methods:

1. If-Then rules: - The knowledge is encoded in the form of condition/conclusion structures. For example, a rule encoding medical information might look like this;

```
IF temperature = 38 and patient = vomiting  
THEN patient = sick.
```

2. Semantic nets: - The relations among objects are represented by links between nodes: e.g. cats are mammals.

3. Frames: - The knowledge about an object is presented in the form of a generalised record structure (i.e. frame format).

4. Horn clauses: - This is a form of predicate logic on which PROLOG (a programming language) is based.

Most systems developed to date make use of the If/Then rule format. Hayes-Roth (1985) says that this is

currently the best way of codifying the problem solving know how of expert systems. While this may be the most widely used method, it must be noted that for some domains the If/Then rule format may not be suitable.

3.2.2. The inference engine

This is the mechanism which the program employs to search and reason in order to solve problem. There are two principle reasoning strategies:

1. Forward chaining: - Here the system works forward from the evidence supplied to it, to the conclusions.
2. Backward chaining: - This is the reverse. Here the system works from the hypothesis and tests for data to either support or refute the hypothesis.

Many systems employ both forward and backward chaining strategies. Buchanan and Duda (1983) describe the reasoning methods as control strategies. Forward chaining is described as data-driven control, while backward chaining is called goal-driven control.

3.2.3. The explanatory interface

A feature of expert systems is that they are able to explain their reasoning methods and justify their conclusions. The explanatory interface provides a partial trace through an expert system's reasoning process, giving easy to understand explanations and reasons on request. It has been suggested that the interface be-

tween an expert system and its human user may be a critical issue in expert system design (Carroll and McKendree, 1987). Coombs and Alty (1984) also emphasize the importance of a system's explanations.

3.2.4. The knowledge-acquisition module

While many expert systems do not contain this component, it has been suggested that expert systems should contain a module which aids in the acquisition of the expert's knowledge. This would assist in the construction of commercial applications by ensuring that the process of prototyping is carried out more easily and more rapidly. In this regard, some researchers have gone as far as to construct separate expert systems which provide assistance in building knowledge-based systems. TEIRESIAS is an example of such a system (Davis, 1979).

3.3. Suitable tasks for expert systems

Various approaches have been adopted to the classification of expert system tasks. Stefik et al. (1982) have identified a number of different types of tasks as being suitable for expert systems to perform. They catalog the following generic expert tasks as possible candidates for expert systems:

1. Interpretation: - Analysing data to determine their meaning.
2. Diagnosis: - Process of fault finding in a system based on the interpretation of data.

3. Monitoring: - Continuously interpreting signals to set off alarms when intervention is required.

4. Prediction: - Forecasting the future from a model of the past and present.

5. Planning: - Creating a program of actions to follow in order to achieve a goal.

6. Design: - Making specifications to create objects that satisfy particular requirements.

Baden (1983) classifies the roles of expert systems from the viewpoint of the end users and identifies the following roles in which expert systems can be used:

1. Consultancy: - Here expert systems are used by non-specialists to obtain specialist advice and help in accomplishing some task.

2. Checklist: - Because expert systems, unlike humans, cannot forget, users can use such systems as checklists. An advantage over a paper checklist is that an expert system could intelligently select or order the questions.

3. Training: - Systems could be used in a training role.

4. Refining expertise: - The system may help in identifying the areas in which the specialists have gaps in their knowledge.

5. Communication medium: - A expert system, with its knowledge base, could be seen to be an active textbook or notebook. Vital knowledge of experts, which is not usually available on paper, could be stored and so made

3. Monitoring: - Continuously interpreting signals to set off alarms when intervention is required.

4. Prediction: - Forecasting the future from a model of the past and present.

5. Planning: - Creating a program of actions to follow in order to achieve a goal.

6. Design: - Making specifications to create objects that satisfy particular requirements.

Baden (1983) classifies the roles of expert systems from the viewpoint of the end users and identifies the following roles in which expert systems can be used:

1. Consultancy: - Here expert systems are used by non-specialists to obtain specialist advice and help in accomplishing some task.

2. Checklist: - Because expert systems, unlike humans, cannot forget, users can use such systems as checklists. An advantage over a paper checklist is that an expert system could intelligently select or order the questions.

3. Training: - Systems could be used in a training role.

4. Refining expertise: - The system may help in identifying the areas in which the specialists have gaps in their knowledge.

5. Communication medium: - A expert system, with its knowledge base, could be seen to be an active textbook or notebook. Vital knowledge of experts, which is not usually available on paper, could be stored and so made

available to others.

3.4. Advantages and Benefits

Expert systems are claimed to have a number of advantages over conventional programs. Basden (1983) lists these as being:

1. Flexibility of expression: - Expert systems are able to embody knowledge in various forms, from rule-of-thumb to more formal expertise. Complex relations can be codified.
2. Human-like processing: - The systems perform a human-like form of reasoning, using rules, facts and levels of relationships. Users can feel more affinity to this type of reasoning. They can also question the system to explain its line of reasoning.
3. Ease of expression: - The language in which the knowledge base is encoded is very close to the language used by the experts. Systems are therefore more easily understood by users who are not necessarily computer literate.
4. Uncertainty: - Uncertain and conflicting data can be handled by expert systems.

The benefits that expert systems hold for commercial applications are revealed by the definition of an expert system i.e. that it solves problems within narrow expert domains. The greatest benefit therefore is that these systems will be able to harness and make full use of an

organisation's scarcest resources namely the talent and experience of key members (Luconi et al., 1986). The author believes that this is the area in which expert systems hold the biggest promise for South Africa, given the scarcity of trained and experienced staff.

Sviokla (1986) makes the point that companies engaged in "expertise intensive" processes who develop expert systems, will gain the most in competitive advantage.

3.5. Disadvantages and Limitations

None of the literature surveyed cites any particular disadvantages which expert systems may have over conventional programs. Banden (1983) makes the point that it is not known for certain what the limitations are since they have not been adequately explored.

Denning (1986) states that the most important limitation of expert systems is their reliability. Expert systems depend on the data in their data bases. They are therefore limited by that information and by the nature of the processes of putting it there. Another point made is that the knowledge elicited has to be encoded. This may result in inconsistent and incomplete knowledge bases which may result in expert systems exhibiting important gaps in knowledge at unexpected times.

Luconi et al. (1986) make the point that knowledge may

be embedded into a system (necessarily incomplete) that is only effective when used by the person who created it. There is also a risk of misapplication when others use the system who lack the knowledge which may be pivotal in the logic leading to a solution.

Finally Denning (1986) states that it is unlikely that expert systems will have clear functional specifications and that their behaviourism in untested situations can be reliably predicted. How easily will users then be able to recognise whether their problem is within the range served by the expert system.

This chapter has briefly examined the concepts of expert systems, their components and the type and nature of the tasks that they are expected to perform. The advantages and limitations have also been discussed. While expert systems are not the magic solution, it can be expected that they will perform useful roles. Bearing in mind their limitations one can set realistic expectations. The following chapter examines examples of expert systems in the domain of income tax.

4. EXISTING INCOME TAX EXPERT SYSTEMS.

This chapter discusses the expert systems whose knowledge domain is income tax, which have been identified in the literature. The following are briefly discussed:

- (a). The particular domain within which the expert system operates.
- (b). The nature of the systems which are employed.
- (c). The tools used.

Issues and problems identified in the development of one such system, ExpertTAX, are also examined. This information was utilised when developing the FOREX system.

Connell (1987), in his review of expert systems applications in accountancy, gives a useful review of income tax based expert systems. Details of the systems identified by Connell are as follows:

1. TAXMAN.

- (a). Domain: The system covers the United States tax rules governing the liability to tax of gains arising from corporate re-organisations.
- (b). System: The relevant facts in each case are presented as a semantic network. A match is then sought.
- (c). Tools: Developed by using micro-PLANNER with some parts being written in LISP.

2. TAXADVISOR

(a).Domain: Financial planning for individuals, namely estate planning in the United States.

(b).System: The knowledge is held in the form of 275 production rules. There is an initial screening process and then evaluations are performed in six consultation subdomains.

(c).Tools: System was built using EMYCIN.

3. CORPTAX

(a).Domain: The system models the law, namely section 302(b) of the United States IR code, which is concerned with stock redemptions.

(b).System: Acts as a checklist, performs calculations, cites and refers to precedents.

(c).Tools: Written in BASIC PLUS.

4. DIRECTORS-TRANSACTIONS

(a).Domain: The system models the law, namely a section of the UK Companies Act dealing with loans to directors.

(b).System: There are 35 IF/THEN rules. Sessions are menu driven and there are good explanation and query facilities.

(c).Tools: Two shells, XI and later Crystal, were used.

5. EXPATAX

(a).Domain: Statutory law, Revenue statements of practice and other material on a narrow but complex area of

UK taxation.

(b).System: The knowledge is encoded in approximately 400 rules. The system is menu driven.

(c).Tools: The system was constructed using the shell APES.

6.TA

(a).Domain: Section 318(a) of the United States IR code dealing with constructive ownership of stock.

(b).System: The system interrogates the user in order to reach a conclusion.

(c). TA is written in PROLOG-86.

7.ACCI

(a).Domain: ACCI deals with the United Kingdom legislation which is used to apportion the income of close companies.

(b).System: The knowledge is held in the form of rules. Sources of information were training notes, statutes, a field manual and an expert.

(c).Tools: The system was built using the shell ADVISOR.

8.ExpertAX

(a).Domain: The tax planning and tax accrual aspects of an audit of an American auditing firm.

(b).System: The system has a knowledge base of over 1000 frames. Frames can be question frames or issue frames. The information was extracted from over twenty different

experts.

(c).Tools: A purpose built shell, Qshell, was used.

9. PAYE

(a).Domain: Rules and legislation concerning the pay as you earn system in the UK.

(b).System: The system models the PAYE rules. Tax documentation was used as the source of knowledge.

(c).Tools: System was written in PROLOG.

The above expert systems are the only ones which have been identified in the literature as being systems in the domain of taxation. Of the systems described, it would appear that ExpertTAX is the most comprehensive and complex system. The system has over 1 000 knowledge frames and it covers both the tax accrual and tax planning aspect of an audit.

Shpilberg and Graham (1987) give a detailed description of the development of ExpertTAX. The issues identified by them during the system's development were the following:

1. Identifying the experts: - Problems arose with identifying who the experts were. The authors concluded that identifying the right expert may be a critical and difficult task when applying expert system technology to some types of problems. They were of the opinion that the more narrowly technical expertise is defined, the more easier it may be to identify the expert.

2. Knowledge base management: - The authors concluded that it remained an open question how the maintenance of expert systems should be managed.

3. System distribution and security: - The point is made by the authors that given the competitive environment, the distribution of the expert system and security aspects become significant issues. Given the fact that an expert system may contain the accumulated knowledge of a firm, organisational and hardware/software controls become important.

4. Other issues: - a). The need to better understand by what is meant by "expertise";

b). More tools and techniques required to evaluate and simulate the real-world environment;

c). Continued research into methods of extracting and capturing expertise.

The only South African expert system in the domain of income tax identified is a system called Tax Partner (Coopers and Lybrand, 1987). The system is described as a expert system for processing tax returns. What the system does is to calculate the income tax liability of all types of taxpayers and prepare the necessary schedules that must be included with income tax returns. The system can therefore be described as a computational expert system and not a advice giving system.

No other income tax based systems have been identified

in the literature. Given the legalistic, rule-based nature of the domain of tax (refer to the next chapter for a description of the domain) it is surprising that not more systems have been developed. One possible explanation for this is that the technology of expert systems is not widely enough known and that the practical problems of applying the technology must still be overcome.

5. THE KNOWLEDGE DOMAIN: SOUTH AFRICAN INCOME TAX.

This chapter describes the knowledge domain in which the expert system was developed. The delineation of the domain and the recognition of the sources of knowledge within the domain is a critical aspect of the initial stage of system development. The chapter describes how the sources of knowledge were identified. The Income Tax Act, No. 58 of 1963, as amended ("the Act") is the primary source of knowledge. The sections of the Act which impact on the domain chosen for the expert system are reviewed. Other expert sources of knowledge such as court cases, text books and expert's writings are examined. The structure and nature of the knowledge is discussed and the chapter ends with an examination of the nature and types of problems for which experts are consulted.

5.1. Identifying Sources of Knowledge

The sources of knowledge first had to be identified before the knowledge acquisition process could commence. The author, being a practitioner in the domain of income tax, is aware of the major sources of knowledge. However when developing FOREX, the author did not want to rely solely on his own knowledge in case this should be incomplete or biased. Other practitioners and experts

(refer chapter 6, section 2) were approached. As part of a general interview they were asked to list the sources of knowledge which they consult when dealing with income tax problems, specifically problems dealing with foreign currency transactions. Answers tended to indicate few sources, with the practitioners indicating that they "knew" the answers to problems.

The structure of the domain and the sources of knowledge identified are discussed below.

5.2. Basis of Taxation

As FOREX is an income tax based system it is useful to briefly examine the provisions of the Act which govern how taxation is calculated.

A popular misconception is that income tax payable is based on profits made. This is not so. The Act lays down a number of steps one must go through without regard to accounting principles. Tax is levied on taxable income received by or accrued to in favour of any person during the year of assessment (section 5(1) of the Act). A number of steps are laid down to arrive at taxable income. One starts with "gross income" which is defined in section 1. All receipts and accruals which are specified as being exempt from tax are then deducted to arrive at "income" (section 1). From this one deducts all amounts which are allowed as deductions in terms of

the Act to arrive at "taxable income".

As can be seen the initial source of knowledge, as codified in the Act, is in the form of a number of steps and rules. This knowledge therefore appears to lend itself to IF/Then rules for the purposes of constructing the expert system.

5.3. Applicable Sections of the Act

A number of sections of the Act apply directly to the taxation implications of foreign currency transactions. These are examined briefly.

Any transactions in foreign currencies which will result in a benefit are brought to account in terms of the definition of gross income. This is the first step as laid down in the Act. Section 1 defines gross income as: "the total amount, in cash or otherwise, received by or accrued to, ... from a source within or deemed to be within the Republic, excluding receipts or accruals of a capital nature,..." .

Similarly, any foreign currency transaction which will result in an expense or loss is governed by section 11(a) of the Act, the general deduction formula. This provides:

"For the purposes of determining the taxable income derived by any person from carrying on any trade within

the Republic, there shall be allowed as a deduction from the income of such a person so derived -
(a) expenditure and losses actually incurred in the Republic in the production of income, provided such expenditure and losses are not of a capital nature."

The only section of the Income Tax Act which applies specifically to foreign exchange transactions is section 24B: "Gains or Losses on Foreign Exchange Transactions". This section provides that profits or losses realised on paying foreign liabilities are taxable or deductible subject to certain conditions being met.

By examining the relevant sections it becomes clear that these can be codified fairly easily into If/Then rules, which is one of the methods of encoding knowledge for expert systems. For example section 11(a) could be restated as follows:

IF expenditure or loss is actually incurred AND
it is incurred in the Republic AND
it is incurred in the production of income AND
it is not of a capital nature AND
it was incurred in carrying on a trade
THEN the amount is allowed as a deduction.

It therefore appears easy to construct an expert system by encoding the relevant sections of the Act.

It was however found on examining the domain in greater depth that, while the provisions of the Act appear straight forward, the interpretation of the facts in each particular case to determine whether the provisions apply, provides more difficulty. For example, applying section 11(a), when is expenditure "actually incurred" or when is it of a "capital" or "revenue" nature?

5.4. Other knowledge sources

The disputes on interpretation have led to many court cases. As the court is the ultimate arbitrator, the decided cases have become the single most important source of expert knowledge. These cases provide much of the knowledge and guidelines for interpreting the facts. It was found when doing the research, that the decisions in various court cases were the crucial factors on which experts based their decisions. When examining the cases which had dealt with transactions in foreign currencies, it was noted that in some cases the judgement dealt with the interpretation of the facts, while in others landmark decisions were made, that is decisions which clarified how the law should be interpreted. For the purposes of extracting knowledge for the expert system, the court cases were found to be an important source of rules, firstly as to what the law is and secondly as how to interpret facts.

Another feature of the knowledge domain noted during the

It was however found on examining the domain in greater depth that, while the provisions of the Act appear straight forward, the interpretation of the facts in each particular case to determine whether the provisions apply, provides more difficulty. For example, applying section 11(a), when is expenditure "actually incurred" or when is it of a "capital" or "revenue" nature?

5.4. Other knowledge sources

The disputes on interpretation have led to many court cases. As the court is the ultimate arbitrator, the decided cases have become the single most important source of expert knowledge. These cases provide much of the knowledge and guidelines for interpreting the facts. It was found when doing the research, that the decisions in various court cases were the crucial factors on which experts based their decisions. When examining the cases which had dealt with transactions in foreign currencies, it was noted that in some cases the judgement dealt with the interpretation of the facts, while in others landmark decisions were made, that is decisions which clarified how the law should be interpreted. For the purposes of extracting knowledge for the expert system, the court cases were found to be an important source of rules, firstly as to what the law is and secondly as how to interpret facts.

Another feature of the knowledge domain noted during the

initial research, is that many of the experts refer to text books as their source of knowledge. It was found that there are two definitive works which were constantly referred to. The two concise, summarized readily available sources of knowledge were found to be an advantage to the author when eliciting the necessary knowledge.

A further examination of the domain showed that, while most problems concerning foreign currency transactions were clear as to the law and only needed interpretation of the facts in a particular case, there were a few areas of dispute and controversy. In these areas, articles written by experts and published in academic journals, were found to be a further source of knowledge. Unfortunately in South Africa this source of knowledge is not very large. It was also noted that, in the areas of uncertainty, experts tend to differ in their opinions.

To summarize, the domain of South African income tax has the following sources of knowledge:

1. The Income Tax Act.
2. Court decisions.
3. Text books.
4. Experts' opinion as evidenced in articles.
5. Experts' opinion as evidenced in discussions.

Before discussing how the necessary knowledge for FOREX was acquired, it is useful to briefly examine the nature and types of problems within the particular domain, for which experts are consulted.

5.5. Nature and types of problems

It is necessary to know the nature and types of problems on which experts are consulted. This is required for determining the type of inference mechanism that is to be used in the system. It is also required for determining what questions the system should ask and what conclusions it must search for.

In the initial questioning of practitioners it was found that the problems concerning the taxation consequences of transactions in foreign currencies could be divided into two categories:

1. Profits or losses made. The majority of problems fall into this category. The question asked is whether profits or losses made as a result of subsequent movements in exchange rates, (whether unrealised or realised) are taxable or deductible. Falling into this category are questions relating to profits and losses on forward exchange contracts.
2. Conversion rates. Here the problem concerns which exchange rate to use for converting foreign currency transactions to S.A. rands. This problem is however often related to the problems under 1.

Having now examined the nature and structure of the knowledge in the domain, the following chapter discusses the techniques used to extract the required knowledge for the FOREX system.

6. THE KNOWLEDGE ACQUISITION PROCESS

Buchanan et al. (1983) state "the process of extracting knowledge from an expert and transferring it to a program is an important and difficult problem". The purpose of this research report, as stated in chapter 2, is to provide practical guidelines for the knowledge elicitation and encoding process. This chapter describes how the necessary knowledge for FOREX's database was elicited. The chapter commences with a brief look at what has been written in the literature on knowledge acquisition and elicitation which was found to be of practical use. Thereafter the process of identifying the experts is described. This is followed by a description of the methods used for eliciting the knowledge from various sources and the sources found to be most useful. The chapter ends with a summary of the methods, the problems encountered and the proposed solutions.

6.1. Literature survey

The knowledge acquisition and transfer processes employed by humans are very complex and a number of disciplines such as psychology, sociology, education, anthropology, philosophy, etc., are involved in their study. It is not the purpose of this paper to discuss or examine the theoretical foundation of knowledge acquisition.

tion. The paper provides practical guidelines and the description of an actual knowledge acquisition and elicitation process undertaken. For this reason only some works, which had some practical use, are discussed.

Gaines (1987) gives some useful background on the theory of knowledge acquisition. In arriving at a operational model of the notion of expertise and the role it plays in society, Gaines examines the phenomenon of expertise acquisition. He identifies a number of problems which may arise when eliciting an expert's knowledge. These are:

1. Expertise may be fortuitous.
2. It may not be available to awareness.
3. It may not be expressible in language.
4. It may not be understandable when expressed in language.
5. Expertise may not be applicable, even when expressed in language.
6. The expertise expressed may be irrelevant.
7. The expertise expressed may be incomplete.
8. The expertise expressed may be incorrect.

Gaines also makes the point that, while most current methodologies for expert systems emphasize interviewing experts, there are also alternative methods for expertise transfer. Examples are; managing the learning environment, evaluation and examples. Gaines concludes his

paper with a knowledge acquisition hierarchy for expert systems in which areas of application for various knowledge acquisition techniques, currently in use, are defined. While the paper provides some background information, it was not of much practical use.

Buchanan et al. (1983) divide the knowledge acquisition process into major stages. These they describe as:

1. Identification stage.
 - 1.1. Participant identification and roles.
 - 1.2. Problem identification.
 - 1.3. Resource identification.
 - 1.4. Goal identification.
2. Conceptualisation stage.
3. Formalisation stage.
4. Implementation stage.
5. Testing stage.
6. Prototype revision.

When developing the system it was found that stages 1 and 2 were of most importance. Planning and proper identification must take place if resources are not to be wasted. Stages 4 to 6 are considered to be beyond the "knowledge acquisition" stage in that this entails the encoding of the knowledge and the construction of the system.

Of further practical use was a "knowledge acquisition

paper with a knowledge acquisition heirarchy for expert systems in which areas of application for various knowledge acquisition techniques, currently in use, are defined. While the paper provides some background information, it was not of much practical use.

Buchanan et al. (1983) divide the knowledge acquisition process into major stages. These they describe as:

1. Identification stage.

1.1. Participant identification and roles.

1.2. Problem identification.

1.3. Resource identification.

1.4. Goal identification.

2. Conceptualisation stage.

3. Formalisation stage.

4. Implementation stage.

5. Testing stage.

6. Prototype revision.

When developing the system it was found that stages 1 and 2 were of most importance. Planning and proper identification must take place if resources are not to be wasted. Stages 4 to 6 are considered to be beyond the "knowledge acquisition" stage in that this entails the encoding of the knowledge and the construction of the system.

Of further practical use was a "knowledge acquisition

grid" developed by LaFrance (1987). This grid was developed to assist in the manual transfer of expertise. The grid describes a two dimensional space. The horizontal plane constitutes five forms of expert knowledge, namely:

1. Layouts:- Map of the task, the boundaries, organisation and basic classifications.
2. Stories:- Classic cases and typical examples.
3. Scripts:- Sequential and procedural knowledge of the domain.
4. Metaphors:- Alternative images of the task, including unique features, constraints and options.
5. Rules-of-thumb:- Tactics and heuristics for interpreting and dealing with circumstances encountered when carrying out the task.

On the vertical plane, LaFrance lists six different types of questions that a knowledge engineer can ask when eliciting knowledge:

1. Grand tour:- Wide questioning in order to understand the boundaries.
2. Categorizing the categories:- Details of the experts terms and concepts.
3. Ascertaining the attributes:- Here the aim is to obtain details of the distinguishing features and range of possible values of the expert's concepts.
4. Determining the interconnections:- These questions seek to uncover the relations among the concepts in

the domain.

5. Seeking advice:- These questions reveal the expert's recommendations and his strategy for how to deal with a variety of conditions.

6 Cross-checking:- These questions validate and examine the limits on previously obtained information.

Some of the above concepts such as the major stages of the knowledge elicitation process (Buchanan et al. (1983)), the forms of expert knowledge and the different types of questions (LaFrance (1987)), were applied when acquiring the necessary knowledge for FOREX. This is discussed below.

6.2. Identifying the experts

The author, along with a number of chartered accountants, is employed by the Department of Finance, Inland Revenue. It is among these chartered accountants, who are involved on a daily basis in the administration of the Income Tax Act, that the initial survey and interviews took place for the purpose of identifying suitable experts.

Initially a number of interviews were held with some of the practitioners. These were selected at random and only a few were interviewed. As only an overview of the domain was initially required, it would have been a waste of time and duplication of effort to interview all

practitioners. The interviews were wide and shallow discussions so that excessive details could be avoided and so that the time spent interviewing could be kept to a minimum. The purpose was to establish:

1. The nature and types of problems that arose within the subject domain.
2. The sources of knowledge consulted by the practitioners when making decisions.

The findings have been reported in chapter 5, sections 4 and 5. These initial surveys are considered to be analogous to the Buchanan et al. (1983) concept of the identification stage or the intersection between Lay-out and Grand tour on LaFrance's (1987) grid.

The major point that came out of the interviews was that, contrary to the authors expectations, the level of knowledge displayed by most practitioners was not very high. This meant that most of those interviewed would not be suitable for extracting the detailed knowledge required for FOREX. The finding highlights the importance of distinguishing between practitioners (those who work in the knowledge domain) and experts (those who have in-depth knowledge of the particular domain). This distinction between practitioner and expert is also dealt with by Hartley (1984).

From the initial interviews two practitioners, who had

in-depth knowledge, were identified as meeting the requirements of the definition of expert.

6.3. Extracting the knowledge

A second, more detailed interview was held with one of the identified experts who appeared to have the most knowledge of the sub-domain in question. The expert in question had previously written a dissertation on the tax implications of foreign currency transactions (Coates, 1987). He was therefore considered to be the expert with the most theoretical knowledge. At this point it can be noted that the author himself has considerable knowledge of the domain of income tax as he works in it and he has previously written an unpublished paper (Els, 1987) on the particular sub-domain. During the initial stages of the research the author did not draw on this knowledge in order to forestall any problems that might arise because of bias, incomplete or incorrect knowledge.

In the interview an in-depth analysis was done of:

1. All the possible types of transactions that could occur in foreign currencies.
2. The taxation implications of the various types of transactions.
3. The different combinations of factors that could have a bearing on the tax consequence.

In the interview it was noted that the expert placed heavy reliance on the theoretical work which he had written. Towards the end of the interview the expert's knowledge was compared to that of the author and the knowledge was effectively combined.

In retrospect it was realised that much of the initial knowledge could have been extracted from text books, law case reports and academic articles (as summarised in the expert's dissertation). This could then subsequently have been verified by checking with the experts. The result would have been a saving of time by the knowledge engineer in the elicitation of the knowledge as well as keeping the use of the expert's expensive time to a minimum.

6.4. Analysis of the knowledge elicited

Using the information extracted and in conjunction with the expert a diagrammatic representation of the knowledge was drawn in the form of a tree diagram. This diagram is reproduced in appendix A.

The tree is a diagrammatic representation of the tax consequences that can occur for the various types of transactions in foreign currencies. This knowledge is what the output of FOREX would be during a consultation i.e. the amount is taxable/deductible etc.

The diagram was constructed by starting at the broadest definition/problem, namely the tax consequences. This was broken down into ever decreasing components. The diagram was found useful for the following reasons:

1. It highlighted the thought process involved in discerning tax implications of foreign currency transactions, indicating which steps the expert follows.
2. It ensured completeness, by providing a systematic, graphical framework in which it was not easy to overlook any important or necessary factor.
3. The flow of the structure laid the foundations for the If/Then rules.

Such similar decomposition of the problem was found to be useful by Ow and Smith (1987) in the development of their expert system.

It should be noted that three different classes of knowledge exist in the domain. These are knowledge concerning:

1. Situations provided for in the Act:- That is rules as to the tax effect of a specific transaction stated unambiguously in the Act.
2. Situations not provided for:- Here experts make up the rules as to the tax consequences of a transaction based on the rules above, court decisions, interpretations and experience.
3. Rules for interpreting facts:- How to interpret the

facts of a case in order to reach a decision under 1 or 2 above.

The knowledge elicited up to this point did not include any in respect of the third type per above. (Refer to section 7. Prototype refinement where the elicitation of this type of knowledge is discussed.)

4.5. Knowledge validation

At this stage it was decided to validate the acquired knowledge. The validation was done in two ways:

1. Live expert confirmation:- The data obtained was discussed with other experts for their confirmation of the tax consequences of a foreign currency transaction given a set of facts.
2. Theoretical confirmation:- Here the knowledge was compared to previous written research done.

In retrospect it was found that written research was the best source of knowledge and the best method of validating knowledge types 1 and 2 (refer section 4 above). This stems from the legal nature of knowledge in the domain.

At this stage it can be noted that in the domain of income tax, expert bias is an important factor to consider when extracting knowledge. The practising tax consultant, as an expert, tends to present only those

arguments which are in his clients favour. Given the nature of the domain, people always seek the most advantageous economic answer for themselves or their clients. Thus neutral, research motivated works were considered to be the most reliable source of knowledge and the most effective way of countering bias.

6.6. Prototype system

At this stage a prototype system was developed to see whether the domain was suitable for further development of the system. The system was found to be feasible. (Refer to chapter 7.)

6.7. Prototype refinement

Refinement of the system commenced and further interviews were held with the experts to elicit their knowledge with regard to how they interpret facts and which rules-of-thumb they employed. (The third information type as discussed in section 4 above.)

The information was obtained by asking direct questions. For example, the distinction between capital and revenue is very important. The expert was asked to list items which he considered to be capital items, why this was so and the questions he would ask to determine such facts. It was found that while the court cases were the source of some such rules, this type of knowledge was

best acquired from the expert in the field. The expert, being involved on a daily basis, has evolved and refined such rules-of-thumb.

It can be noted that the refinement and further development of the system is a process which proceeds hand-in-hand with further knowledge acquisition. This was found to be the easiest and most effective route. As problems concerning the knowledge cropped up during the encoding process, the experts would be consulted.

6.8. Findings.

6.8.1. Identification of experts

During the initial stages of the knowledge elicitation process, the identification of the experts was found to be a problem. It is important to make the distinction between practitioners and experts. While most practitioners have a working knowledge, very few were found to have sufficient in-depth knowledge of the more complex and difficult areas in the domain. The identification of the expert is therefore a critical aspect of the initial stage of the knowledge elicitation process. The use of the incorrect "expert" may result in the system having faulty knowledge, which in a domain of a legal nature such as income tax, could render the system useless. The solution is to ensure that the persons selected meet the requirements of being experts and to validate the knowledge extracted against some neutral source.

6.8.2. Knowledge elicitation

The methods used and found to be the most effective in acquiring the knowledge for FOREX's database in the domain of income tax were:

1. Examination of written knowledge sources.
2. Interviewing of experts.

For the particular domain of income tax it was found that a substantial amount of the knowledge could be found in written works such as text books, law case reports and articles. This was especially relevant for information concerning the tax consequences of transactions. Indeed, these sources were found to be the ones consulted by experts themselves when making their decisions.

For knowledge concerning how to interpret the facts of a situation, the interviewing of experts was found to be the best way of extracting the knowledge. While it is true that much of this knowledge is available in the reports of court cases, it is more efficient to extract this from the experts who make daily use of such knowledge in their work.

6.8.3. Countering bias

A further problem encountered is that of bias in the knowledge. This is especially relevant in a domain such as income tax. Firstly there is the problem of bias in

6.8.2. Knowledge elicitation

The methods used and found to be the most effective in acquiring the knowledge for FOREX's database in the domain of income tax were:

1. Examination of written knowledge sources.
2. Interviewing of experts.

For the particular domain of income tax it was found that a substantial amount of the knowledge could be found in written works such as text books, law case reports and articles. This was especially relevant for information concerning the tax consequences of transactions. Indeed, these sources were found to be the ones consulted by experts themselves when making their decisions.

For knowledge concerning how to interpret the facts of a situation, the interviewing of experts was found to be the best way of extracting the knowledge. While it is true that much of this knowledge is available in the reports of court cases, it is more efficient to extract this from the experts who make daily use of such knowledge in their work.

6.8.3. Countering bias

A further problem encountered is that of bias in the knowledge. This is especially relevant in a domain such as income tax. Firstly there is the problem of bias in

the system developer's knowledge, if he is a practitioner in the domain. If the developer places too much reliance on his own knowledge the system may display incomplete or incorrect knowledge. The solution developed to counter this bias was not to rely at all initially on the developer's own knowledge and subsequently, when using such own knowledge, to validate this against that of other experts.

Secondly, in the specific domain of income tax, experts' knowledge may be biased in that only favourable answers are sought. More validation of the knowledge to written works was found to be an effective counter to any possible bias.

6.8.4. Domain expertise of knowledge engineer

The required knowledge for FOREX was extracted fairly easily from the experts (once these had been identified) and from written works. By being a practitioner in the domain, the author was able to interact much more easily with the experts and was able to understand the knowledge acquired. Little time needed to be spent on obtaining an understanding of the problems, terms and concepts of the domain. It was found that having a knowledge of the domain was a definite advantage when acquiring the necessary knowledge.

the system developer's knowledge, if he is a practitioner in the domain. If the developer places too much reliance on his own knowledge the system may display incomplete or incorrect knowledge. The solution developed to counter this bias was not to rely at all initially on the developer's own knowledge and subsequently, when using such own knowledge, to validate this against that of other experts.

Secondly, in the specific domain of income tax, experts' knowledge may be biased in that only favourable answers are given. Here validation of the knowledge to written works was found to be an effective counter to any possible bias.

6.8.4. Domain expertise of knowledge engineer

The required knowledge for FOREX was extracted fairly easily from the experts (once these had been identified) and from written works. By being a practitioner in the domain, the author was able to interact much more easily with the experts and was able to understand the knowledge acquired. Little time needed to be spent on obtaining an understanding of the problems, terms and concepts of the domain. It was found that having a knowledge of the domain was a definite advantage when acquiring the necessary knowledge.

5.8.5. Further research

A point to make however, is that the knowledge acquisition and elicitation process is fairly lengthy. If the time which the experts spent researching their theoretical work is taken into account, a substantial amount of time was required to acquire the knowledge, which was within a very narrow domain. This problem is a bottleneck in the development of expert systems and it will prevent the the development of commercial applications. This is identified as an area for further research.

7. ENCODING THE KNOWLEDGE.

As stated in chapter 3, there are four different ways of representing knowledge in an expert system. This chapter describes which of the methods of encoding the acquired knowledge was used. The problems of controlling search strategy and knowledge organisation encountered are discussed and the solution developed for use in FOREX is described.

Dhar and Pople (1987) make the point that "the major decision for a expert system builder is how the expert knowledge is to be represented". Knowledge can be represented in four different ways (refer chapter 4):

1. IF/THEN rules.
2. Semantic nets.
3. Frames.
4. Horn clauses.

Ramsey et al. (1986) identify the three most important methods of building expert systems to be bayesian classification, rule-based deduction and frame-based deduction. They discuss the relative strengths and weaknesses of the different approaches as well as the situations in which each method is easy or difficult to use.

During the initial stages of the research, the author intuitively settled on the rule-based method of

knowledge representation. As the research progressed the method selected was confirmed as being suitable. As discussed in chapter 5 the knowledge structure is rule-based. The primary source of knowledge, the Income Tax Act, is in effect a rule book; i.e. If x, y and z then amount b is allowed as a deduction.

As discussed in chapter 6 section 3 a distinction can be drawn between two types of knowledge in FOREX's domain. Firstly there is the information concerning the tax consequences. Some consequences are spelt out explicitly in the Act. These are the easiest to code into If/Then rules, and such encoding was achieved in a short space of time. For example a part of section 24B (dealing with realised losses on foreign liabilities) was encoded in the following rule form:

```
IF amount = realised AND
   type = capital OR type = revenue AND
   difference = loss AND
   Sec24B_cond1 = yes
THEN tax_effect = deductible
```

Similarly, situations where the tax consequences are not provided for specifically in the Act, but where the experts opinion was taken, the information was easily encoded into rule form.

The second type of information, dealing with how the

facts of a case should be interpreted for the purposes of determining which situational rule applies, was found also to be easily encodable into rule form. For example a rule concerning the distinction between revenue and capital items was encoded as follows:

```
IF underlying_transaction = debtor OR
   underlying_transaction = stock OR
   underlying_transaction = cash OR
   underlying_transaction = working_capital
THEN type = revenue.
```

It is concluded that the actual process of encoding the acquired knowledge into IF/THEN rule format, presented no problems. It was however discovered that the rules could not be programed into the expert system shell in a haphazard manner. Cognizance had to be taken of the search strategy employed by the inference engine of the shell. [In this case VP-Expert (Sawyer et al. (1987)), refer chapter 8.]

It was noted during the development of the prototype that the system would not work properly until it was programmed how and what to search for.

It was found that creating a control structure and setting the problem solving activity was far more difficult and important than coding the knowledge into rules. The literature survey shows that substantial research is

taking place in this area. For example, Davis (1980) considers the problems of reasoning about control and suggests the concept of Meta-rules. Georgeff (1982) proposes a general production system architecture that allows procedural control knowledge to be directly presented and used. The concept of using domain specific knowledge to support opportunistic reasoning is discussed by OW and Smith (1987). They also propose an hierarchical organisation structure to control and co-ordinate problem solving activity.

This latter concept was also applied in the development of FOREX. During the interviews with the experts it was noted that they, when considering a problem, asked certain questions first. For example, when considering profits or losses on transactions, the first line of enquiry was whether such transactions were of an asset or a liability nature and whether the profits or losses were realised or unrealised. This concept of decomposing the problem was discussed in chapter 6.

This concept has also been applied to the organisational structure and procedural control system of FOREX. It was found that this approach had two advantages:

1. The problem solving activity was more easily controlled in that the system was channelled into appropriate paths to follow.
2. The development and maintenance of the system was

made much easier by compartmentalising the knowledge base. Each module could be developed independently and tested separately. Additions to the modules could also be done easily by adding submodules as necessary.

8. UTILITY OF USING AN OFF-THE-SHELF SHELL.

8.1. Shell selection

Expert systems can be programmed in four different ways (Forsyth, 1986):

1. Shells
2. AI environments
3. AI languages
4. Conventional languages

As stated in chapter 2, this research does not address the problems of hardware and software selection. The research was undertaken from the point of view of a non-specialist in expert systems. It was therefore decided to make use of an expert system shell. The shell VP-Expert was selected for use. The system was selected purely on the basis of availability. No comparisons with other shells were done nor were the specifications of VP-Expert used as a basis of selection. Those seeking further information on the different expert system shells available are referred to Gevarter (1987), who discusses in detail various expert system building tools in terms of their capabilities.

8.2. VP-Expert

VP-Expert (Sawyer et al. (1987)) is a cheap, widely

available expert system shell. The shell is a rule based system which employs a backward chaining inference engine. The shell comes on a single floppy disk and can be used on any IBM or IBM compatible personal computer. The system comes with a easy to read manual.

The way VP-Expert works, is that the system is set a goal. This is to search for the value of a variable. The inference engine then searches the knowledge base for a rule naming the variable in its THEN conclusion. It then seeks to match data to the variables and conditions contained in the IF portion of the rule. This data is obtained by either finding the values of other variables named in the THEN conclusions of other rules or by asking questions. Once values of all the variables and conditions have been found and these match those contained in the IF portion of the rule, the rule is said to be passed and the value of the goal variable is accordingly assigned. This backward chaining strategy can be effectively employed in the domain of income tax.

8.3. Experience with VP-Expert

The author was able to acquaint himself with the system within a few hours. A basic prototype (based on 14 rules) was constructed within approximately two hours. The development of FORSX, while taking some thirty hours, was accomplished easily by using the shell.

No major problems were found when using VP-Expert other than the weakness of the explanatory facility. As discussed in chapter 3, one of the advantages of expert systems is that they can explain why they are asking a question and how the conclusion was arrived at. This interface facility is very important in a consulting role (Corroll and McKendree, 1987). The facility, while available in the shell, could not be used effectively. VP-Expert provides that a user can ask the question "how?" in order to find out why the system has given a particular answer. When this question is asked by a user, all the variables for which values were obtained during the consultation are listed. The user is then asked which variable he is enquiring about. Should the user not be the developer, he will have great difficulty in selecting the right variable as these would be unknown to him. A partial solution is to keep the names of the variables as close as possible to their true meaning. In FOREX, the user is told at the beginning of the consultation to enquire about the variable "tax_effect" when asking the question "how?". The system will then give the reasons for why the answer was arrived at. As no comparison was done with other shells, it is not known whether the same problem exists in other systems.

3.4. Advantages of using a shell

The advantages of using expert system shells as

No major problems were found when using VP-Expert other than the weakness of the explanatory facility. As discussed in chapter 3, one of the advantages of expert systems is that they can explain why they are asking a question and how the conclusion was arrived at. This interface facility is very important in a consulting role (Carroll and McKendree, 1987). The facility, while available in the shell, could not be used effectively. VP-Expert provides that a user can ask the question "how?" in order to find out why the system has given a particular answer. When this question is asked by a user, all the variables for which values were obtained during the consultation are listed. The user is then asked which variable he is enquiring about. Should the user not be the developer, he will have great difficulty in selecting the right variable as these would be unknown to him. A partial solution is to keep the names of the variables as close as possible to their true meaning. In FOREX, the user is told at the beginning of the consultation to enquire about the variable "tax_effect" when asking the question "how?". The system will then give the reasons for why the answer was arrived at. As no comparison was done with other shells, it is not known whether the same problem exists in other systems.

§.4 Advantages of using a shell

The advantages of using expert system shells as

Identified by the author are:

1. Availability of expert systems to laymen:- Such shells make available the technology of expert systems to a wide range of users who are not specialists in the domain of expert systems.
2. Ease of use:- Anybody with only some basic computer literacy (such as the author) will find the shells easy to use.
3. Cost:- Some shells, such as VP-Expert, are very cheap thus making available the technology to a wide range of users.
4. Time saving:- Prototype systems can be developed within very short periods of time. Working models within very narrow domains, such as FOREX, can be developed in a few manweeks.

It can be concluded that developing FOREX using the VP-Expert shell was an easy task for a layman in expert systems technology. The development was done within a period of approximately thirty hours. The area of weakness identified concerns the interface between the system and the human user.

9. VALIDATING THE SYSTEM.

One of the most important criticisms against expert systems concerns their reliability (Denning, 1986). Refer to chapter 3.5 . The validation of a system is therefore a important component of the construction task. This chapter explains the methods employed to validate FOREX and discusses the results of the validation tests.

9.1. Approaches to validation

Buchanan and Duda (1983) state that the results of empirical experiments are still the best indicators of validity. They state that these may best be accomplished by randomized studies and double blind experiments. They point out that logic analysis of completeness and consistency will be inadequate for establishing the validity of an expert system. However static checks can reveal potential problems.

O'Keefe et al. (1987) in a comprehensive descriptive article, discuss the problems of expert system validation and present both qualitative and quantitative methods of validating expert systems. They state that the major problems in evaluating expert system performance are:

1. What to validate.

2. What to validate against.
3. What to validate with.
4. When to validate.
5. How to control the costs of validation.
6. How to control bias.
7. How to cope with multiple results.

The qualitative methods of validation (which are subjective comparisons of performance) identified by them are:

1. Face validation: - Subjectively comparing the system performance against human expert performance.
2. Predictive validation: - Using historic test cases to compare performance.
3. Turing tests: - Blind evaluation by comparing human experts' decisions (without the experts knowing) to the decisions of the system.
4. Field tests: - Testing the system in the field, noting performance errors when they occur.
5. Subsystem validation: - The system is decomposed into subsystems and each subsystem is validated one at a time as they are developed.
6. Sensitivity analysis: - Systematically changing input variables and observing the effect upon the system performance.
7. Visual interaction: - Visual animation of the system workings is followed by the experts.

The quantitative validation techniques, which are statistical tests which compare the performance against test cases or human experts, consist of:

1. Paired t-tests.
2. Hotellings One Sample T-square test.
3. Simultaneous confidence intervals.
4. Consistency measures.

The author, not having sufficient theoretical knowledge of statistical methods, found that the methods described above were of no practical use.

O'Leary (1987) has developed a validation framework with reference to systems in auditing and accounting applications. The contents of the framework are:

1. Content validity: - Representativeness of the content which is validated by direct examination of the system and by testing the system against humans and other models.
2. Criterion validity: - This refers to the criteria used to validate the system such as comparing system decisions with expert decisions. The level of expertise of the system must be defined by some set of criteria. The knowledge base criteria include consistency, accuracy and completeness.
3. Construct validity: - What is meant here is that a tried and tested theory should exist on which the system is based.
4. Objectivity: - When validating a system any bias and

variance should be minimized. This could be achieved by having programmer validation (if the programmer has no vested interest), by having independent administration of validation, by end user validation, by using blinding techniques and by using different development and test data.

5.Economics: - Cost-Benefit analysis.

6.Reliability: - The system should generate identical solutions to identical inputs.

7.Systematic (experimental) variance: - The test problems used must be designed in such a way that one can distinguish between systematic variance and chance. The problems should reflect the range of problems, there must be sufficient variation in the problems and there must be a sufficient number of problems.

8.Extraneous variance: - This should be controlled where possible. Factors such as system complexity, the systems position on its life cycle trajectory, location of judges during testing (in the laboratory or in the field) and blinding during validation effect extraneous variance.

As briefly discussed above, the topic of validation is a fairly wide and comprehensive one. It is not the objective of this report to discuss the merits of the various validation techniques. This chapter describes the techniques used to validate FOREX.

9.2. Validating FOREX

As described in chapter 8, the system was developed in stages by compartmentalising the knowledge base. As each different section was completed, random tests of each section were carried out by testing whether the system gave the correct solution to various types of problems. (The subsystem approach using face validation and predictive validation as described by O'Keefe et al. (1987).) These test case problems only tested the validity of the conclusion. The system's interpretation of facts was not tested. The initial tests were based on the author's expertise in the domain. These early tests showed that logic errors easily crept into the system's knowledge base and highlighted the need for thorough testing.

Once the system was completed the experts who had previously been interviewed during the knowledge acquisition process were approached. They were requested to develop a number of hypothetical problems on which an expert could be consulted. These they then tested on FOREX. The solution given by the system was compared to the experts' solution. The experts were also asked to comment on the way the system worked and the questions the system asked for the purpose of modifying the user interface if required. Once these tests were completed modifications, where necessary, were made to FOREX.

These tests also showed that validation is an important aspect of the construction process. A few errors were still detected in the system.

At this stage the final version of FOREX was completed. It was now decided to test the system against actual cases. As the author works at Inland Revenue, he has access to a wide range of actual cases. A list was obtained of cases on which Inland Revenue officials had come across problems concerning the tax implications of transactions in foreign currencies. These cases were tested on FOREX by using the facts of the case as stated in written correspondence contained in the files. The solutions proposed by FOREX were compared to the decisions made by the officials concerned. The results of the tests showed that in most cases the solution proposed by FOREX was the same as the decision made by the officials. In those cases where the answers were not the same, the difference could be attributed to the fact that the experts also differ at times as to what the correct solution to a problem is. This is especially true in those areas where the Income Tax Act does not specifically prescribe and where no court cases exist.

Given the very narrow domain and small knowledge base of FOREX it was not very difficult to validate the system. The nature of the domain also assisted in the validation process. In the domain of tax there are a limited number

of conclusions. For example, an amount is either taxable or not taxable or it is deductible or not deductible. No other possibilities exist. It is therefore concluded that the validation of any income tax based expert system should not present problems with regard to the completeness of the knowledge base. Problems may however arise as to correctness, especially in those areas where uncertainty and disagreement as to the correct tax consequence exists among experts. In the case of FOREX these problems only effect a very small percentage of transaction types. This should not effect the expert system's utility. The systems should be designed in such a way that when an answer is given in an area of uncertainty, the system mentions the uncertainty, the alternative solution and the most likely correct answer. This feature has been incorporated into FOREX in the area where disagreement among experts exists.

The following chapter will present a sample consultation using the completed system.

and rate of exchange. By implication no adjustment is necessary if the loss relates to expenditure. The effect on taxable income is the same whether the expense is shown at the correct value or if a loss on exchange is claimed as a deduction. If the liability relates to stock an adjustment will be necessary to the value of closing stock to the extent that closing stock is on hand. Refer to Caltex Oil v S.I.R.

The unrealised revenue loss of 20 100 is deductible CNF 100.

Press any key to continue."

It can be noted that the system has given the answer (second last line) and has expanded by providing a further explanation. The answer also restates the facts of the particular case; i.e. "unrealised" "revenue" "loss" which is quantified. The CNF 100 means a confidence factor in the answer of 100%.

At this stage the user wants to know how the answer was arrived at. He enters [/Zhow]. The system responds with:

"Which variable are you enquiring about?

The variable "tax effect" is selected. The system shows:

"The loss is deductible because the court held in the case of Caltex Oil v SIR that expenditure must be accounted for at the end of the year of assessment. Therefore any expenditure in foreign currencies is quantified by using the rate of exchange ruling at the end of the year of assessment."

The consultation is continued by pressing any key.

"Do you wish to have a further consultation?
yes no "

Enter "yes".

"Does the next consultation concern:
1). Taxability/deductibility of transaction in
foreign liabilities OR
2). Other transactions. 2 "

The user selects "2".

The system breaks for a few seconds while it chains to the original beginning. [Refer to the begin of the consultation documented.] As can be seen the system ask two types of questions. The one type controls the search direction and operations of the system, while the other type requests the information on which the system makes its decision.

The system responds when it is asked why it is asked a question as well as how it arrived at a answer. The latter facility is however not very understandable as the system asks the user which variable he is enquiring about. Should the user not know how the VP-Expert shell operates it is doubtful whether he will obtain an answer. This weakness of the explanatory interface has been commented on previously (refer chapter 8).

11. CONCLUDING FINDINGS

This research report has documented the development of a income tax based expert system. The system, named FOREX, is in the domain of the taxation consequences of transactions in foreign currencies. The system was developed with the purpose of providing practical guidelines for others who wish to develop their own expert systems not necessarily in the tax domain. The research addressed:

1. The knowledge acquisition process.
2. The encoding of the knowledge.
3. The utility of using a expert system shell.

The paper first examined the concept of expert systems. The definition and features of a expert system were discussed. The components of a typical system were examined. The tasks which expert systems perform, the advantages and the limitations of such systems were also discussed.

11.1 Tax-based expert systems

Various expert systems in the domain of income tax, which had been identified in the literature, were discussed. It was noted that most systems have very narrow domains and that not many systems have been developed to date. It is suggested that this may be because the technology of expert systems is not yet widely enough

known and that there are still practical problems in applying the technology.

11.2. The domain of income tax

The paper then examined the intended domain for FOREX, namely South African income tax with specific reference to transactions in foreign currencies. The sources of knowledge were discussed as were the nature and types of problems on which experts were consulted. It was found that important sources included written works such as text books, law case reports and articles as well as human experts. It was also found that the knowledge in the domain (as contained for example in the Income Tax Act) was easily encodable into If/Then rule format.

11.3. Knowledge acquisition and elicitation

11.3.1. Expert identification

The initial stage of the knowledge acquisition process is the identification of the real experts. It was found that this presented some problems. The level of expertise of all available practitioners was not at the required level. The identification of proper experts was found to be a crucial first step of the knowledge acquisition and elicitation process.

11.3.2. Knowledge elicitation techniques

The interviewing technique was used to acquire much of the required knowledge. It was however found that in the

domain of tax, much of the initial knowledge could be extracted from written work such as text books, law case reports and written articles. The technique of decomposing the knowledge into subproblems was found to be very useful. This proved useful when checking completeness of the knowledge as well as later, when encoding the knowledge.

11.3.3. Validation of elicited knowledge

It can be recommended that the acquired knowledge be validated before encoding commences. This will preclude incorrect knowledge from being included in the data-base of the system. Confirmation of the elicited information with experts and testing against written theoretical works were found to be effective techniques of validating the elicited knowledge.

In the domain of tax the knowledge is often biased, in that the experts seek to obtain the most favourable tax effect. In this case the knowledge was validated by obtaining expert and theoretical confirmation.

11.3.4. Practitioners as system developers

It was found that it is an advantage if the system developer is a practitioner in the domain. This aids in the understanding of the domain specific terms as well as in ensuring the completeness and accuracy of the acquired knowledge.

11.4. Encoding of acquired knowledge

The acquired knowledge was encoded easily into If/Then rule format. This part of the knowledge encoding process presented no problems. It was however found that it is important to create a proper control structure for the problem solving process. The solution developed was to use a hierarchical organisation structure along the lines of the tree diagram developed during the knowledge acquisition process.

11.5. Utility of using a shell

Using an off-the-shelf shell presented no problems. The shell is easy to use and the system development was accomplished fairly quickly. It can be concluded that using a shell in system development is a viable route for laymen in expert system technology to follow.

11.6. Validating the system

FOREX was validated in three stages. Firstly the developer performed random tests to confirm the accuracy of the end conclusions while developing the system. The system was then tested against hypothetical cases prepared by the domain experts. The experts also performed the tests so that the user interface could be tested. Once all adjustments had been finalised, the system was tested by the developer against actual cases contained in the files of Inland Revenue.

11.6. Summary

This research has shown that it is fairly easy for a layman in expert system technology to develop a system within his domain of expertise using an expert system shell. The availability of expert system shells makes the technology easily accessible. The knowledge in this case could be extracted from written works and interviewing techniques could be employed. This particular mix might not be appropriate in other situations. Identifying the real experts may present some problems. An area for future research which has been identified is the lengthy process of acquiring the knowledge. This is a bottleneck in the development of expert systems. Finally experience in this research supports the recommendation that expert systems be developed for very narrow domains. This makes the development easier. The knowledge that is to be acquired is kept to manageable levels, the encoding process is made easier, the system control structure can be kept simple and validation is made more certain.

11.7. Further possibilities

The author envisages that expert systems can be constructed for the whole domain of income tax by constructing numerous autonomous systems in various narrow sub-domains with ultimately a separate system to control the selection of the appropriate system for a consultation.

REFERENCES

Baden, A., 1983: On the Application of Expert Systems. International Journal of Man-Machine Studies, Vol. 19, pp 461 - 477.

Bobrow, D.G., Mittel, S. and Stefik, M.J., 1986: Expert Systems: Perils and Promise. Communications of the ACM, Vol. 29, No. 9, pp 880 - 894.

Buchanan, B.G., Borstow, R.B., Bechtal, R., Bennet, J., Clancey, W., Kulikowski, C., Mitchell, T. and Waterman, D.A., 1984: Constructing an Expert System. Building Expert Systems, Hayes-Roth, F., Waterman, D.A. and Lenat, D.B. (eds), Addison-Wesley, London, pp 127 - 167.

Buchanan, B.G. and Duda, R.O., 1983: Principles of Rule-Based Expert Systems. Advances in Computers, Yovitz, M.C. (ed), Vol. 22, pp 163 - 216.

Carroll, J.M. and McKendree, J., 1987: Interface Design Issues for Advice-Giving Expert Systems. Communications of the ACM, Vol. 30, No. 1, pp 14 - 31.

Coates, A.E., 1987: The Tax Implications of Foreign Exchange Profits and Losses, Unpublished Paper submitted for the Higher Diploma in Income Tax, University of the Witwatersrand, Johannesburg.

Computer Mail, 1988: Realism at Last, Vol. 3, No. 7, pp 59 - 62.

Coombs, M. and Alty, J., 1984: Expert Systems: an Alternative Paradigm, International Journal of Man-Machine Studies, Vol. 20, pp 21 - 43.

Coopers and Lybrand, 1987: Tax Computation Software, TeX Flash, July, No. 3.

Davis, R., 1980: Meta-Rules: Reasoning about Control, Artificial Intelligence, Vol. 15, pp 179 - 222.

Davis, R. and Shortliffe, E., 1977: Production Rules as a Representation for a Knowledge-Based Consultation Program, Artificial Intelligence, Vol. 8, pp 15 - 45.

Dhar, V. and Pople, H.E., 1987: Rule-Based v Structure-Based Models for Explaining and Generating Expert Behaviour, Communications of the ACM, Vol. 30, No. 5.

Ela, D.U., 1987: The Taxation Treatment of Foreign Exchange Transactions, Unpublished Paper submitted for the Degree Master of Commerce (by Coursework), University of the Witwatersrand, Johannesburg.

Forsyth, R., 1986: The Anatomy of Expert Systems, Artificial Intelligence, Yazdani, M. (ed.), Chapman-Hall

Gaines, E.R., 1987: An Overview of Knowledge-Acquisition and Transfer, International Journal of Man-Machine Studies, Vol. 26, pp 453 - 472.

Georgeff, M.P., 1982: Procedural Control in Production Systems, Artificial Intelligence, Vol. 18, pp 175 - 201

Gevarter, W.S., 1987: The Nature and Evaluation of Commercial Expert System Building Tools, Computer, May, pp 24 - 41.

Goguen, J.A., Weiner, J.L. and Linde, C., 1983: Reasoning and Natural Explanation, International Journal of Man-Machine Studies, Vol. 19, pp 521 - 559.

Goul, M. and Tonge, F., 1987: Project IPMA: Applying Decision Support Design Principles to Building Expert Systems, Decision Sciences, Vol. 18, pp 448 - 467.

Hayes-Roth, F., 1985: Rule-Based Systems, Communications of the ACM, Vol. 28, No. 9, pp 92. - 932.

Hartley, R.T., 1984: CRIB: Computer Fault-finding through Knowledge Engineering, Computer, March, pp 76 - 83.

Kelodner, J.L., 1983: Towards an Understanding of the Role of Experience in the Evolution from Novice to Expert, International Journal of Man-Machine Studies, Vol. 19, pp 497 - 518.

LaFrance, M., 1987: The Knowledge Acquisition Grid: a Method for Training Knowledge Engineers, International Journal of Man-Machine studies, Vol. 26, pp 245 - 255.

Lin, E., 1986: Expert Systems for Business Applications: Potentials and Limitations, Journal of System Management Vol. 37.

Luconi, F.L., Malone, T.W. and Scott Morton, M.S., 1986: Expert Systems: the Next Challenge for Managers, Sloan Management Review, Summer, pp 3 - 13.

Mar ins, G.R., 1984: The Overselling of Expert Systems, Datamation, November.

McDermott, J., 1982: XSEL, a Computer Sales Person's Assistant. Machine Intelligence 10, Hayes, J.R. , Hitchie, D. and Pao, Y.H. (eds.), Ellis Horwood, Chichester.

O'Keefe, R.M., Balci, O. and Smith, E.P., 1987: Validating Expert System Performance, IEEE Expert, Winter, pp 81 - 90.

O'Leary, D.E., 1987: Validation of Expert Systems with Application to Auditing and Accounting Expert Systems, Decision Sciences, Vol. 18, pp 468 - 485.

Ow, P.S. and Smith, S.F., 1987: Two Design Principles for Knowledge Based Systems, Decision Sciences, Vol. 18, pp 430 -447.

Pollitzer, E. and Jenkins, J., 1985: Expert Knowledge, Expert Systems and Commercial Interests. Omega International Journal of Management Science, Vol. 13, No. 5, pp 407 - 418.

Ramsey, C.L., Reggia, J.A., Nau, D.S. and Ferrentino, A., 1986: A Comparative Analysis of Methods for Expert Systems, International Journal of Man-Machine Studies,

Vol. 26, pp 435 - 452.

Riesbeck, C.K., 1984: Knowledge Re-Organisation and Style, International Journal of Man-Machine Studies, Vol. 20, pp 45 - 61.

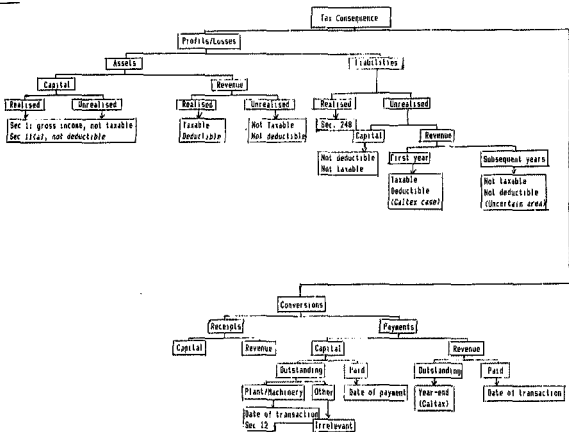
Sawyer, B., Moose, A. and Shafer, D., 1987: VP-Expert, Paperback Software International, Berkley.

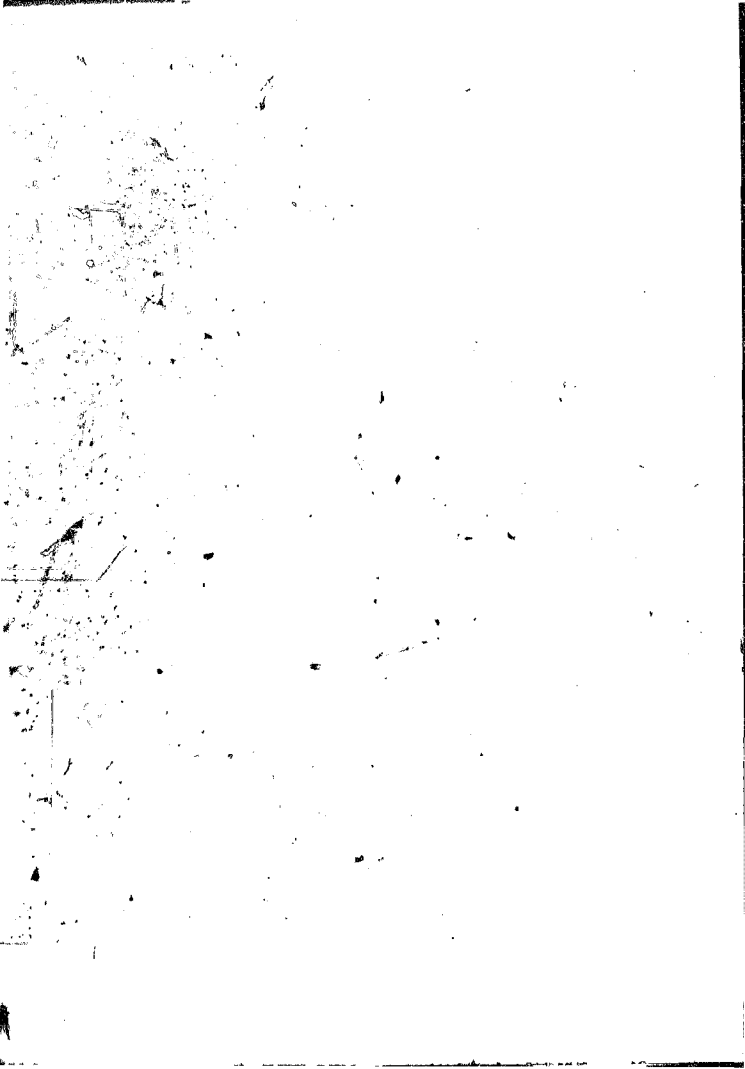
Shpilberg, D. and Graham, L.E., 1986: Developing Expertax: an Expert System for Corporate Tax Accrual and Planning, Auditing: A Journal of Practice and Theory, Vol. 6, No. 1, pp 75 - 94.

Stefik, M., Alkins, J., Balzar, R., Benoit, J., Birnbaum, L., Hayes-Roth, F. and Sacerdoti, E., 1982: The Organisation of Expert Systems: a Tutorial, Artificial Intelligence, Vol. 18, pp 135 - 173.

Swiokla, J.J., 1986: Business Implications of Knowledge Based Systems, Data Base, Summer, pp 5 - 19.

APPENDIX A





Author Els Dirk Udo

Name of thesis Using An Off-the-shelf Shell To Develop An Income Tax Expert System. 1988

PUBLISHER:

University of the Witwatersrand, Johannesburg

©2013

LEGAL NOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed, or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page) for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.